AD NUMBER
AD804600
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Critical Technology; NOV 1966. Other requests shall be referred to Aerospace Medical Research Laboratories, Attn: MRHTO, Wright-Patterson AFB, OH 45433. This document contains export-controlled technical data.**
AUTHORITY
AFHRL ltr, 5 Nov 1970

AMRL-TR-66-196

AN ASSESSMENT OF RESEARCH RELEVANT TO PILOT TRAINING

ALFRED F. SMODE
EUGENE R. HALL
BIOTECHNOLOGY, INC.

DONALD E. MEYER
AEROSPACE MEDICAL RESEARCH LABORATORIES

NOVEMBER 1966

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of Aerospace Medical Research Laboratories (MRHTO), Wright-Patterson Air Force Base, Ohio 45433.

AEROSPACE MEDICAL RESEARCH LABORATORIES
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

NOTICES

600 - December 1966 - C0192-17-327

BLANK PAGE

AN ASSESSMENT OF RESEARCH RELEVANT TO PILOT TRAINING

ALFRED F. SMODE
EUGENE R. HALL
DONALD E. MEYER

FOREWORD

This report represents a part of the study effort in a program to organize research findings and information relevant to pilot training. The work was performed under contract AF 33 (615)-2968 between the Behavioral Sciences Laboratory of the Aerospace Medical Research Laboratories, and BioTechnology, Inc. Dr. Alfred F. Smode was the Principal Investigator for the study which was carried out during the period 1 February 1966 to 30 June 1966.

The research was performed under Project 1710, "Human Factors in the Design of Training Systems," and Task 171003, "Human Factors in the Design of Systems for Operator Training and Evaluation." Dr. Gordon A. Eckstrand, chief of the Training Research Division, was the Project Scientist and Dr. W. Dean Chiles, Operator Training Branch, was the Task Scientist.

Appreciation is extended to James F. Parker, Jr. for preparing a portion of the text dealing with a review of research on perceptual-motor performance.

We are also indebted to the many scientists who have contributed information and data for this study. These include: J. Roger Berkshire, Robert Blanchard, Hugh Bowen, Wiley Boyle, Paul Caro, W. D. Chiles, Ralph Flexman, Charles Kelley, J. D. Lyons, William McClelland, Wallace Prophet, and Ben Schohan.

This technical report has been reviewed and is approved.

WALTER F. GRETHER, PhD
Technical Director
Behavioral Sciences Laboratory
Aerospace Medical Research Laboratories

ABSTRACT

This report presents a critical review and interpretation of the considerable amount of research data that have either direct or indirect implications for the training of pilots. The purpose is to organize systematically the research findings from the human performance and the training research literature that are pertinent to pilot training, and, based on the status of research in defined areas, to identify researchable issues. Successive portions of the report deal with studies on the definition of the pilot's job, the acquisition of flying skills, performance measurement, simulation and transfer of training, operational components of the pilot's job, and the maintenance of flying proficiency. In addition, attention is given to studies concerned with improving training systems and recent innovations in training methods are reviewed. As it provides a considerable background of information directly concerned with pilot training, this report will be of interest to individuals involved in any aspect of flight training.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

TABLE OF CONTENTS--Continued

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS--Continued

BLANK PAGE

SECTION I

INTRODUCTION

A survey of advanced pilot training programs in the Air Force[1] has revealed some areas in which existing research data could be applied to an advantage, as well as areas in need of further research. A problem exists however, in that much of the pertinent research is difficult to find and "piece together" in a fashion valuable for pilot training. It is on this basis that the research literature relevant to pilot training is being examined.

PURPOSE

This study presents a compilation of resource information applicable to Air Force pilot training programs. It is based on a review and interpretation of the available literature in terms of applicability and usefulness to pilot training programs. The purpose is twofold: (1) to provide a state-of-the-art survey of research findings from the human performance and the training research literature that is truly useful for pilot training, and (2) based on the status of research in defined areas, to identify researchable issues for improving portions of the pilot training program.

PERSPECTIVE

The present effort centers on the question, "What resources does the technical literature provide that are relevant to improving pilot training; e.g., what is currently known from over two decades of research that can be translated effectively into pilot training programs?" The question is a complicated one since it is difficult to define precisely what areas of a vast body of literature pertain to the sophisticated job of military flying, and further, to determine the extent to which data in any area are relevant to the complex world of operational training. Although for many years people have deplored the seeming irrelevance of the available research data for pilot training, we were unable to locate any

---

[1] Based on a survey of advanced pilot training requirements in the Air Force (Smode & Meyer, 1966; Smode, Post, & Meyer, 1966). This study describes Air Force training requirements for representative combat and logistics aircraft/mission combinations and the practices currently employed in producing the operationally ready pilot.

study whose purpose was to review the research literature and determine its applicability to pilot training. What is available are reviews of portions of the literature in areas subsumed under pilot training or somewhat relevant to pilot activities.

The work began with a search of an imposing number of study titles listed in the technical abstracts of the Defense Documentation Center (DDC, located in Alexandria, Virginia). The technical abstract bulletins and their indexes were surveyed for the period January 1960 to May 1966, and a demand bibliography relevant to pilot training and pilot performance was prepared by DDC which yielded about 150 titles. Documents were also selected based on several dozen relevant descriptors, the most rewarding of which were: pilots, performance, transfer of training, training devices and simulators, education, tracking, human engineering, and operator personnel. In addition, emphasis was placed on locating available studies prepared in programs of research on pilot performance by such organizations as: the Civil Aeronautics Administration, the Air Force Human Resources Research Center, the Personnel and Training Research Center, the Human Resources Research Office of the George Washington University (Army Aviation), and the Naval School of Aviation Medicine. Finally, the standard sources of psychological and behavioral science literature were consulted. All told, several hundred prime studies were so identified. Conferences were also held with individuals knowledgeable in aviation training.

Because of the mass of literature, that by titles alone appeared eligible for review, the more recent studies were selected. Research published from 1960 was most actively sought. This decision was based on the awareness that the greatest proliferation in training technology has occurred since that time. However, this is not the whole story. It was necessary to go back further in time because various programs of research on pilot training were accomplished from 1950 to 1960, and a sizable number of studies were conducted in the period bracketing World War II. In fact, "surges" in published research are related to the existence of groups that conducted research specifically on pilot training. The only available systematic study efforts have come from these organizations, who--as part of their research program--have attempted to resolve the significant problems in pilot training. The number of correlated studies in the literature corresponds with the peak time periods these units were actively conducting this type of research. Table I lists these important centers for pilot training research, their time periods of operation, and their major research emphasis.

TABLE I

CENTERS OF PILOT TRAINING RESEARCH

Unit	Time Period of Operation	Research Emphasis	Location
Civil Aeronautics Administration	Pre-WW II, latter 1930s	Training publications, methods, and performance measures.	Washington, D. C.
USAAF Aviation Psychology Program	WW II	Pilot selection techniques, task analyses, proficiency measures, training methods, selection of instructors.	Ft. Worth, Texas
USAF Human Resources Research Center (HRRC) (pilot training laboratory)	Oct 1949 to Jan 1954	Pilot training requirements, simulator research and evaluation, pilot inflight evaluation, instructor methods.	Randolph AFB, Tex.
USAF Personnel and Training Research Center (AFPTRC) (pilot performance laboratory)	Feb 1954 to Apr 1958	Continuation of HRRC research.	Eglin AFB, Fla.
Human Resources Research Office, Army Aviation Research Division	Dec 1958 to present	Training, motivation, leadership, and training device requirements for Army aviation related jobs.	Ft. Rucker, Ala.
National Aeronautics and Space Administration	Oct 1958 to present	R&D test pilot selection and training, astronaut selection and training, performance measures, and basic research.	Langley AFB, Va. Edwards AFB, Calif. Houston, Texas Moffett Field, Calif.
Federal Aviation Agency	Early 1950s to present	Task analysis, training publications, performance measures, basic research.	Washington, D. C. Atlantic City, N.J. Oklahoma City, Okla.

A substantial number of studies available in the open literature that deal with some aspect of human performance were identified for review. These researches covered a considerable range of content areas, employed diverse methods, and were set in a variety of conceivable laboratory, simulation, or field contexts. Since results of these studies varied in applicability to pilot training programs, a foremost concern was the determination of what was applicable to our purpose. To set a realistic theme, a number of ground rules were established. These are set forth below and provide the rationale for the review.

1. The review centers on training research studies dealing with pilot and pilotlike (e.g., aerial observer, automobile driver) activities. Studies providing human performance data relevant to the pilot's job are included, but emphasis is placed on the performance data that permit a better understanding of the critical aspects of training--for example, information that can be used knowledgeably in improving training efficiency. Thus, the event(s) must be modifiable by means of training in order to be considered in the review. Excluded from the review are studies requiring a major extrapolation on our part to bring the findings in line with the business of training pilots. Many laboratory studies, for example, may appear relevant but cannot be reconciled with pilot training requirements. Studies of this sort may be cited but are not treated in a primary way since it is not within anyone's capability to judge their relevance to pilot training. To attempt extrapolation would be doing a disservice to the present program.

2. An emphasis is placed on research for training (i.e., meshing the pilot with the aircraft); studies for the purpose of equipment design are omitted.

3. Studies involving perceptual-motor behavior and inflight activities are emphasized in view of the specific requirements of the pilot's job.

4. Researches yielding generalizable data are sought; therefore, evaluation studies which compare a specific component or a specific system against some standard are omitted.

5. The review considers only those studies pertinent to the pilot's job as defined by present and immediate future aircraft requirements. Knowledgeable Air Force officers agree that pilot training requirements for the present and for the immediate future are, for the most part, similar (for example, transitioning from the F-105 to the F-111 aircraft poses no greater set of difficulties than transitioning from the F-100 to the F-105 aircraft).

The research studies appraised in this report are of two basic kinds: (1) studies conducted specifically in the aviation context where the results pertain directly to aspects of flying and pilot training, and (2) studies of human performance in more general or nonaviation contexts where the results are relevant to pilot training (our decision) and hence, are described as having utility for pilot training.

MAGNITUDE OF USAF PILOT TRAINING

Aside from the operational advantages resulting from the use of efficient training techniques, improvements in training have significant economic value because of the size of the military pilot training effort and the length of time required for training. At present, eight U.S. Air Force bases are engaged in fixed-wing undergraduate pilot training designed for an output of approximately 3,000 pilots per year. The training program is 53 weeks in length. This undergraduate pilot training is estimated to cost about $77,000 per pilot. To this figure must be added the costs of the combat crew training programs conducted within the major commands. Although no average cost figure is available, it is estimated that the undergraduate training costs are more than doubled by the time a combat-capable pilot is produced. In a keynote address to a conference on engineering systems for education and training, The Honorable Thomas D. Morris, Assistant Secretary of Defense (Manpower),[2] asserted that pilot training is the most costly and time consuming in-house training effort within the military establishment. Excluding depreciation of facilities and investment in aircraft, the training of a jet pilot costs about $250,000, while $110,000 and $45,000 are needed respectively to train a propeller aircraft pilot and a helicopter pilot. The annual cost for this training approaches one billion dollars. It is clear that even small gains in pilot training efficiency would result in substantial savings in dollars per year.

ORGANIZATION OF THE REPORT

In order to arrange and discuss the wide content range selected for this report, the following distinctions were made. All research was

---

[2] Keynote address by Thomas D. Morris, Assistant Secretary of Defense (Manpower), to a Joint Department of Defense, Office of Education, and National Security Industrial Organization Conference on Engineering Systems for Education and Training, Marriott Motor Hotel, Arlington, Virginia, 14-15 June 1966.

considered that dealt with the analysis and definition of the pilot's job, including the behavioral requirements underlying performance. Also of prime importance was the considerable body of research describing training and performance effects in pilot activities. An emphasis was placed on the acquisition process in pilot skill development, flight simulation and transfer of training, performance measurement, and the maintenance of pilot proficiency. Finally, a class of research was considered that dealt with aspects of the training system for structuring pilot training programs.

Accordingly, the content of this study is organized into three major sections, in addition to the Introductory Section. Section II deals with the description and definition of the pilot's job. Section III reviews the data from general laboratory research and from training research dealing with the acquisition, assessment, and retention of flying skills and the behaviors pertinent to the pilot's job. Section IV presents findings of research for improving training systems which are particularly relevant to the purpose of this review. In each section a number of research areas are defined and, within each, the pertinent literature is reviewed. Some areas are more extensively treated than others. This is the result of differences in the amount of pertinent published literature or the importance of the area for pilot training and research. Following the review, research issues are suggested for continuing effort within the defined and/or allied areas. The philosophy underlying the research recommendations is that each issue outlined is either (1) specific to an obvious need in pilot training programs, or (2) intended to increase the understanding of human behavior which may prove fruitful in solving training problems not easily defined.

In each section, researchable issues with forseeable utility are developed based on the status of research in the areas reviewed. These point up the need for (1) more data on human performance in tasks and situations highly representative of the pilot's job under conditions found in flying, and (2) studies which define training effects as a function of differing methods, media, and schedules applicable to the training system. The need for validation studies is underscored. For the most part, emphasis is placed on programmatic research and laboratory and simulation studies (supported by field data) not necessarily geared to quick response in problem solution. In an earlier report (Smode & Meyer, 1966) based on a survey of Air Force Combat Crew Training Schools (or the equivalent) and operational units, research was proposed that could yield reasonably immediate and useful results for improving training within the types of units visited. These issues had the following features: were oriented to field needs (quick response and implementation), were

meaningful in the operational context, and had high face validity (readily perceived as an issue of importance). Although the emphases in the present study and in the earlier study are quite different, a certain amount of overlap exists between the two sets of research issues.

The recommendations to be made here are based on one or more of the following:

1. More data are needed along the current defined lines of inquiry.

2. Better organization and application of known data are needed.

3. Reorientation of thinking on the problem/area is required. Present approaches and solutions represent "dead ends," and new and perhaps "radical" approaches are needed to enrich theory and enable formulation and quantification of currently difficult variables.

An interesting sidelight in this review is that some problems which received attention in the literature at an earlier time have diminished in significance today. Advances in design technology (improvements in design, unburdening the pilot through the use of automatic assists), new procedures, etc., have minimized problems which formerly were substantial. Some of the more prominent examples follow.

Shifting of research emphasis away from manual control problems to greater concern for information-processing, decision-making, and time-sharing behavior in complex man-machine systems.

Improvements in aircraft instrumentation and in-flight displays such as developed by the Army-Navy Instrumentation Program (ANIP) and Joint Army-Navy Aircraft Instrumentation Research (JANAIR) programs (e.g., contact analog) diminish the appropriateness of the dichotomy of instrument versus contact flight.

Shift in emphasis of crew coordination research due to the trend toward smaller crew complements in aircraft and changing mission requirements.

Advances in digital simulation techniques which have improved the fidelity of aircraft simulators.

Changing emphasis in vigilance research due to changes in the structure of monitoring tasks and monitoring requirements.

Diminution of the personnel subsystem time lag in system development.

Changing emphasis in the training of components of flight (e.g., greater emphasis on survival in hostile environments, on tactics training, and less on aerobatics, etc.).

Thus, some groups of studies earnestly pursued at an earlier time may suggest problems that are less valid in the current context, and for this reason, are not emphasized as researchable issues.

PROBLEMS IN TRANSLATING THE LITERATURE FOR USE IN PILOT TRAINING

The general conclusion from the analysis and interpretation of the literature cited is that very little of the results is directly applicable to pilot training. The body of the findings simply does not contain the substance needed for resolving major problems in pilot training. Perhaps the basic reason for this has been the absence of systematic or programmatic assaults on the prevalent issues to be resolved. The major research needs have been known and expressed in one form or another for many years, but the research has not been structured in terms of these needs. Most often, there has been a sporadic "chipping away" at portions of the defined issues with no overall concepts of guidance enunciated by users, buyers, or researchers.

The inability to apply much of the results of research to pilot training is explainable in several more specific ways. In particular, training research which presumably strived to obtain end products of specific value in the training process for the purpose of enhancing training efficiency and human performance, has not been effectively directed at improving pilot training, as evidenced by lack of implementation of results. What has been studied has been dictated by the availability of apparatus and equipment and by task situations relatively easy to install. Important (and sticky) issues have been conveniently bypassed in going this route. There are, of course, exceptions (for example, the research on parameters involved in low-altitude, high-speed flight), but by and large, "convenient" experimental task situations have predominated.

There is also substantial ambiguity surrounding the research that deals specifically with pilot performance. Various shortcomings can be identified in the experimental procedures and tasks. The more telling of these are outlined below.

Noncomparability of measures across studies (e.g., different measures of proficiency used, such as accidents, attrition rates, nonstandard flight checks, ratings on different inflight events).

Heavy reliance on subjective opinions. Instructor ratings on "goodness" of performance are the most available and center on what the individual instructor considers important. It is difficult to know what constitutes the elements of criterion performance. Thus, differing bases for comparisons exist and the results of a study become highly specific to that study.

Differing tolerance limits for describing adequacy of performance during inflight measurement (e.g., differences in out-of-tolerance envelope).

Procedural changes within a study as it progresses; for example, subjects may be transferred, equipments modified or changed during the study, scheduling and administrative problems may occur, and more rarely, changes may be dictated because of safety considerations. The result is a severely unbalanced design.

Differing ways of interpreting transfer-of-training data. In some cases, transfer assessments may be based on performance in initial trials; in other cases it may be based on performance across larger blocks of the transfer task.

Reporting of the same studies in several different documents, making it difficult to determine exactly what was done.

Use of imprecise criterion measures of the event being examined. The criterion is sometimes irrelevant or confounded in assessing the effects of the independent variable.

Differences in skill level of pilots/trainees, making for noncomparability among subjects.

These conditions which are, for the most part, peculiar to research on flight performance, compromise seriously the capability to compare the results of similar studies so that in many instances only approximations or guesses can be made about the status of events in a research area. Certainly, the difficulties in conducting systematic experimental research on pilot training are underscored by these shortcomings.

Additional variance is contributed by the difficulty of relating research tasks to the pilot's task in flying an airplane. This is part of the more general and traditional problem of correlating laboratory and simulation conditions with real world conditions. This turns out to be especially true of the complex activities which define the pilot's job. The laboratory tasks are simplified abstractions of the real job. Generalizing from this is done with significant risks. An inability to define precisely the pilot's job must be added to the problem. The classic dichotomy used to describe pilot tasks has been in terms of continuous skills and procedural skills. Tracking behavior, which is represented in the manual control of an aircraft, is a case in point. Much research has been aimed specifically at understanding this type of skillful response which involves a continuous interaction of input, output, and feedback processes. In the simple tracking example, the operator is provided a visual display which presents an error signal. He manipulates a control mechanism which generates an output signal in response to the display change, which results in the reduction of the error signal. The requirement is to null the error signal; thus, graded response is required in continuous interaction with stimulus changes. Engineering psychologists have, with scientific vigor, employed physical servomechanism theory as the model for the man-machine tracking system. Yet, the impressive amount of research to date has pointed up the difficulties in achieving a coherent theory of tracking behavior. The inability of servoanalysis to handle nonlinearities in human behavior, and the inability to interrelate higher conceptual processes (mediating responses, motivation, etc.) with the independent variables which influence tracking behavior and measures of performance are the cogent reasons for the difficulty in theory development. Thus, one's enthusiasm for generalizing from this data base to the control of an aircraft is somewhat curtailed. And, in practical terms, tracking behavior (although complex) is but a part of the whole array of sequential and interactive behaviors that a pilot must accomplish during flight.

In addition to the weakness in the power to generalize from the research findings, studies employing complicated equipments to provide the task structure (e.g., simulators) are often too comprehensive to achieve meaningful conclusions. Complex experimental designs (used to make full use of costly experimental equipment) and equipment malfunctions unfortunately introduce inconsistencies into the data. The effects of the variables become obscured and the results are often confounded and difficult to interpret.

In essence, not much of the research information is immediately useful for pilot training in terms of the stringent requirements assumed above. The data are weak for the reasons cited. The research nevertheless yields a variety of suggestions that are worthy of use and refinement in training. Although few unequivocal results exist, the body of data provides the necessary clues for zeroing-in on the prevalent training problems and providing information for delimiting the range of variations predictable for performance. The data also serve to suggest more precise and meaningful hypotheses for continuing research. While the review makes explicit the weaknesses in the research, it tacitly assumes the above features.

As an aside, it appears that the most immediate, pressing problem is one of achieving a rapprochement between laboratory research and the needs of the operational environment. The dilemma is this: a significant amount of research has been accomplished presumably to aid in the accomplishment of real-world events. But the people in the operational environment are forced to solve their immediate problems as best they can (often an inadequate or costly measure) because of lack of research-supported data. Much of the research is unorganized and unsystematic by any standards and is of indeterminate value to pilot training. One reason already cited is the "piecemeal" nature of the research, done by individuals who are not organized to coordinate their efforts with other researchers.

The shortcomings in research support for pilot training pose an interesting question: "What is an effective way out of the dilemma?" Probably the most obvious answer is the need for a research effort having as its minimum requirements the following: a group made up of scientific and operational personnel (research teams) to initiate and monitor needed programs; an emphasis on "on-site" research and application; a capability for longitudinal studies, as required; and an emphasis on obtaining validity data in a training program. Such an "organization for training," responsive to changing field requirements, would also provide documentation procedures and media to take advantage of previous

experience and avoid duplication of effort. Achieving this is easier said than done. Individuals close to the problem, nevertheless, approve of such a proposal as a means necessary to make full use of the existing training technology and to put into perspective research needs and required research effort.

DESCRIPTION OF THE PILOT'S JOB

The job of the pilot is the "given" in this study, the commodity with which we are dealing. Some idea of the difficult behavioral requirements of the job and environment in which the Air Force pilot performs is outlined here as a prelude to the review and evaluation of specific areas of the literature pertinent to pilot training. The demands and complexities of the pilot's job make research on pilot training an elusive undertaking.

Pilot activities are performed with the body in a seated or set position, and skilled performance is based on the ability to handle an array of information from separate sources and organize various stimuli or inputs. This is somewhat different from skilled performance with the body in motion (e.g., athletes, soldiers) which involves large-amplitude motion and force components. Where the individual and the primary external object are in motion simultaneously, (e.g., a quarterback throwing a football to a moving receiver), swiftly changing relationships are in interplay which are not found in performance in the seated position. The responses of the pilot include relatively simple procedural or discrete acts (e.g., positioning levers, switches, and controls; communicating verbally; etc.) and continuous manual control movements requiring small forces and a sensitivity to pressure exerted on the controls. A premium is placed on the integration of responses, coordination and timing, time-sharing, decision-making, and judgmental processes. Fitts (1962) has identified four crucial aspects of skilled tasks performed by the pilot:

> Cognitive (task understanding, knowledge of equipment, strategy, judgment, decision making, planning).
>
> Perceptual (what to look for, identifying relevant cues, making critical discriminations of forces, pressures, dividing attention).
>
> Coordination (integration of movements, timing of successive movement patterns).

Tension-relaxation (anxiety, motivation components, skill development in effort and timing movements).

An additional feature, not overly emphasized, is that a pilot is required to store a considerable amount of information of many classes both short-term (emphasizing immediate memory requirements as found in navigation, penetration and landing, air-to-ground communication, etc.) and long-term (emphasizing flight rules, procedures, tactics, etc).

In its simplest expression, flying has been dichotomized into procedural responses (discrete mediating responses) and continuous control responses (multidimensional tracking responses based on contact or flight instrument cues). Detailed attempts at describing the components in the pilot's job have been set forth by a number of investigators (for example, Gagne, 1962; Fitts, 1962; Smode, Gruber, & Ely, 1962). In essence, the defined components include: procedural acts (discrete motor, verbal, mediating responses); perceptual-discriminative acts (differentiating stimulus aspects, identification and monitoring, anticipation responses); perceptual-motor acts (graded response in continuous interaction with stimulus changes); and decision-making, concept-using acts (selection of alternative behaviors and use of operating rules which may include the processing of complex information prior to initiating a course of action). Of course, categories of this sort introduce a degree of artificiality since discrete beginnings and endings are conveniently identified for behaviors which, in reality, are not so easily defined. However, separate tasks in flying are identifiable in such a gross classification scheme embracing such important pilot requirements as, controlling the movement of the aircraft about its three axes, accomplishing "housekeeping" and ancillary tasks, monitoring and time sharing, and making decisions and programming for the immediate future based on processed multiple information.

The complexity in flying comes however, not from the definable task demands per se but when these are constrained by the forced-pace nature of performance emphasizing time-sharing requirements. The acceptable performance envelope requires that sequential and coordinative activities be correctly accomplished within a specified time frame (i.e., skills performed under time stress for speed and accuracy of response). The pilot is accountable for a number of events simultaneously. He must keep track of many separate sources of information and stimuli, collate these, and sort out effects produced by his own earlier actions from the effects produced by outside events. Some idea of the totality of the demands on the pilot can be grasped from the following additional requirements imposed by the operational setting.

Communications/navigation (hearing, recalling, remembering, and initiating actions when called for, usually under conditions of high volume ground-air communications).

Sensitivity to aircraft characteristics

Weather, turbulence, icing, poor visibility

Emergencies, malfunctions, aircraft degradation

Air traffic in terminal areas, terrain

Flight rules, airspace constraints

Crew interaction

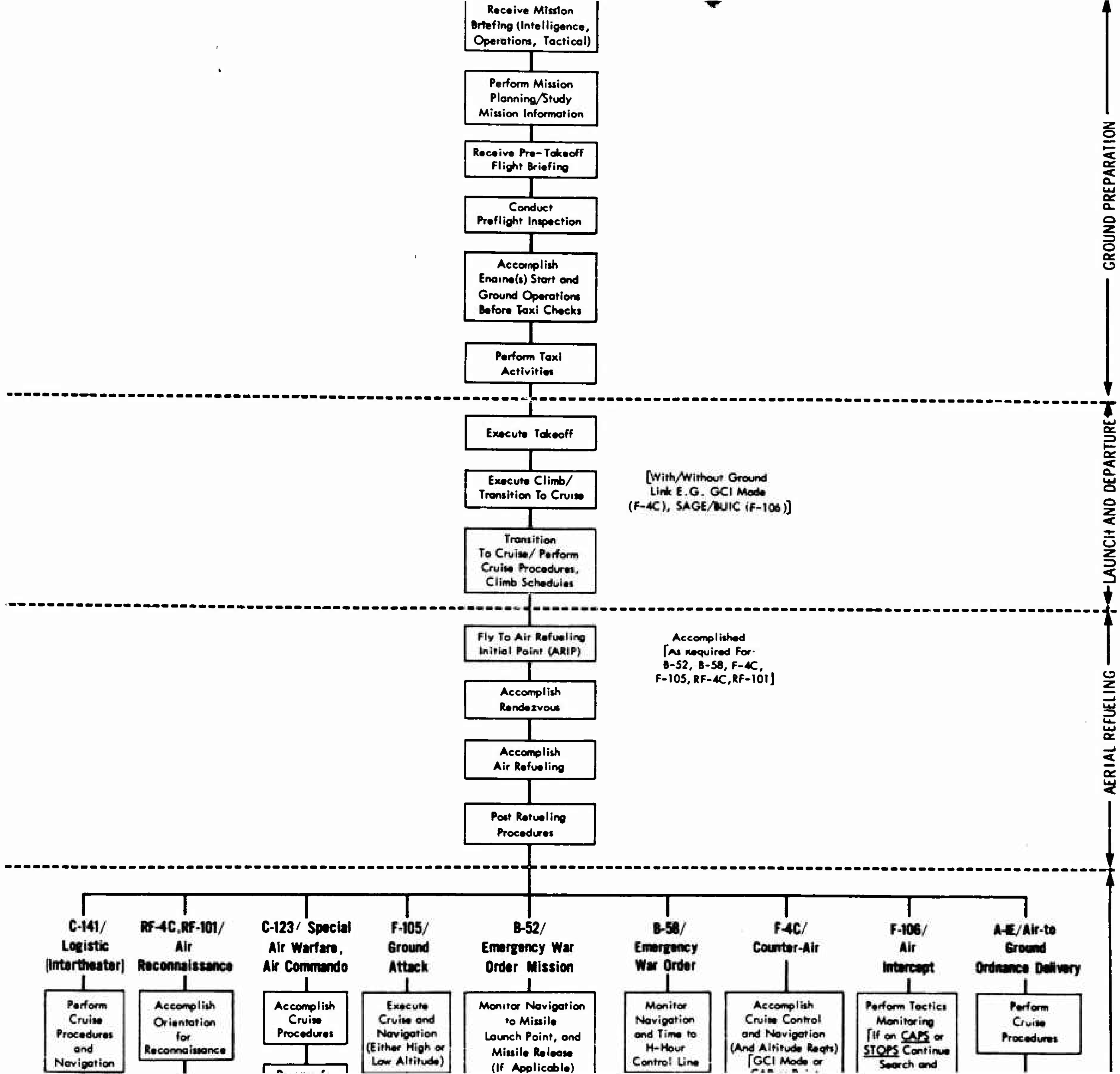
Mission planning, anticipation of events during flight

The consensus is that flying is a complex job requiring a number of different abilities. This was stated clearly by Miller (1947) in summarizing the experiences of the Army Air Force Pilot Project during World War II:

> The fact that the best printed tests in the classification battery used during the war were found to predict flying ability at least as well as the best apparatus tests, confirms the observation that perception, visualization of spatial relationships, knowledge of mechanical principles, and motivation are important factors in learning to fly. The fact that the addition of apparatus tests to this battery improved its ability to predict, suggests that motor skill and ability to perform a complex paced task are also important factors.

Exactly what these ability factors are and the relative importance of each has yet to be determined.

The behavioral requirements involved in flying receive additional temporal anchoring when placed in a specific mission profile for any of the defined Air Force missions. An indication of the complexity of flight operations is shown in Figure 1. This presents a composite of the flight


GROUND PREPARATION
LAUNCH AND DEPARTURE
AERIAL REFUELING
Receive Mission Briefing (Intelligence, Operations, Tactical)
Perform Mission Planning/Study Mission Information
Receive Pre-Takeoff Flight Briefing
Conduct Preflight Inspection
Accomplish Engine(s) Start and Ground Operations Before Taxi Checks
Perform Taxi Activities
Execute Takeoff
Execute Climb/ Transition To Cruise
[With/Without Ground Link E.G. GCI Mode (F-4C), SAGE/BUIC (F-106)]
Transition To Cruise/ Perform Cruise Procedures, Climb Schedules
Fly To Air Refueling Initial Point (ARIP)
Accomplished [As required For: B-52, B-58, F-4C, F-105, RF-4C, RF-101]
Accomplish Rendezvous
Accomplish Air Refueling
Post Refueling Procedures
C-141/ Logistic (Intertheater)
Perform Cruise Procedures and Navigation
RF-4C, RF-101/ Air Reconnaissance
Accomplish Orientation for Reconnaissance
C-123/ Special Air Warfare, Air Commando
Accomplish Cruise Procedures
F-105/ Ground Attack
Execute Cruise and Navigation (Either High or Low Altitude)
B-52/ Emergency War Order Mission
Monitor Navigation to Missile Launch Point, and Missile Release (If Applicable)
B-58/ Emergency War Order
Monitor Navigation and Time to H-Hour Control Line
F-4C/ Counter-Air
Accomplish Cruise Control and Navigation (And Altitude Reqts) [GCI Mode or
F-106/ Air Intercept
Perform Tactics Monitoring [If on CAPS or STOPS Continue Search and
A-1E/Air-to Ground Ordnance Delivery
Perform Cruise Procedures

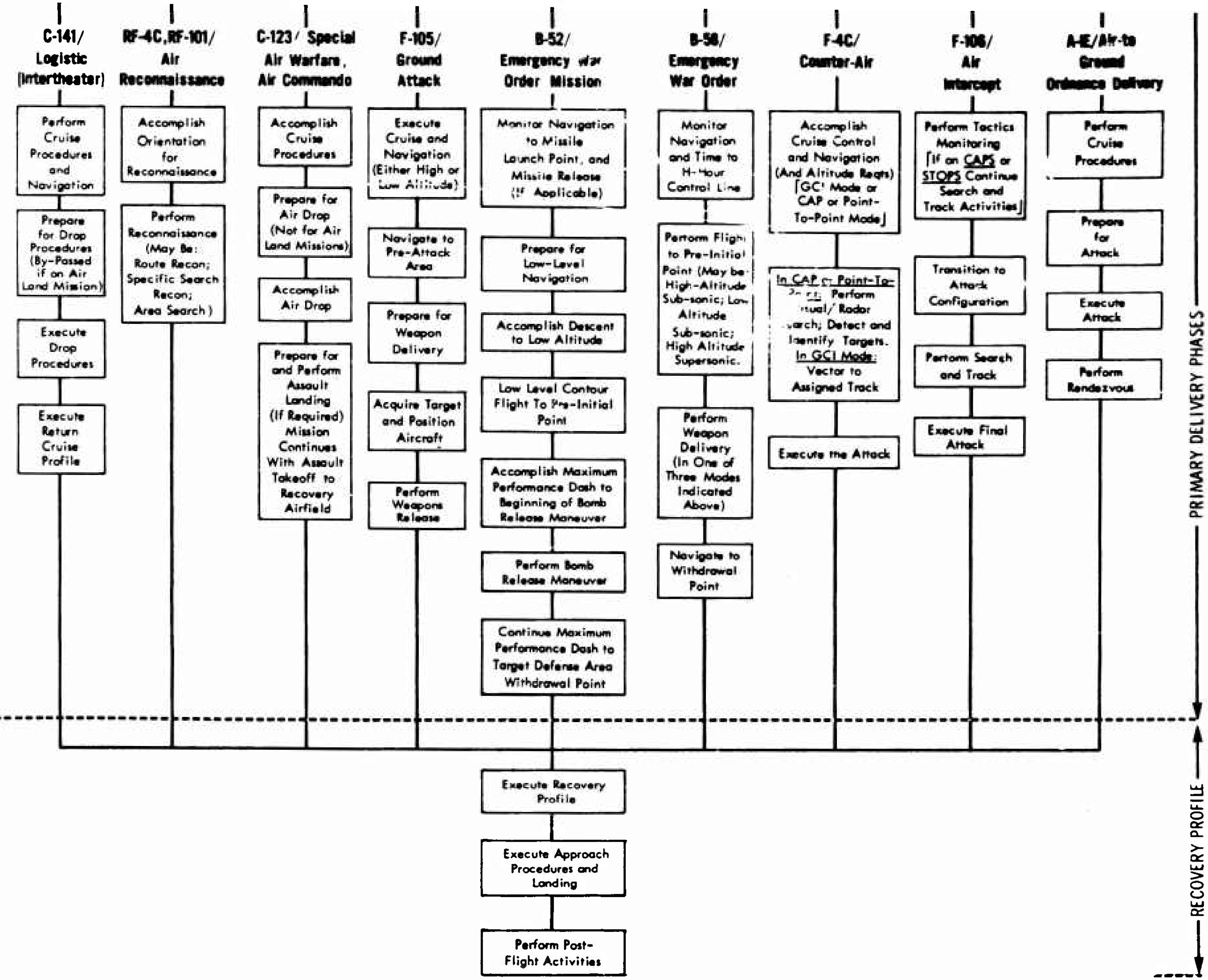


Figure 1. Composite of Flight Segments and Tasks Involved in Current Air Force Combat/Logistics Missions.

segments and tasks involved in Air Force combat/logistics missions. The data are based on nine of the most primary aircraft/mission combinations current within the four major Air Force commands.[3]

---

[3]Described in an earlier phase of the present research program (Smode, Post, & Meyer, 1966). The aircraft/mission combinations analyzed were: B-58, Emergency War Order (EWO); B-52, Follow-On EWO; F-4C, Counterair (Air Superiority); F-105, Ground Attack; RF-4C/ RF-101, Tactical Air Reconnaissance; A-1E, Air-to-Ground Ordnance Delivery; C-123, Air Commando; C-141, Logistics, and F-106, Air Intercept.

BLANK PAGE

SECTION II

STUDIES DEFINING THE PILOT'S JOB

It is well recognized that flying an airplane is a complex operation for the human and that the extensive knowledge requirements and the dimensions of piloting skills are not clearly understood. An understanding of this complex of tasks, performed under forced-pace and time-shared conditions, and requiring precision, timing, coordination, and heightened channel capacity of the human, has been sought, essentially, via three methodological approaches: experimental, rational, and correlational. It is the purpose of this section to review studies describing and defining aspects of the pilot's job from the viewpoints peculiar to these methods.

The first group of studies considers the laboratory research on perceptual-motor performance. The majority of these researches pertinent to pilot behavior deal with some aspect of manual control; hence, tracking behavior has been a major preoccupation of these investigations.

The second group centers on rational-analytical studies investigating the behavioral composition of piloting tasks in the mission environment.

The last group deals with studies employing factor-analysis methods in definition of pilot job components.

The specific research areas reviewed in this section are organized as follows:

Experimental Research in Perceptual-Motor Performance
Rational Analyses
Correlational Studies

EXPERIMENTAL RESEARCH IN PERCEPTUAL-MOTOR PERFORMANCE

The basic activity of controlling an aircraft requires a skilled perceptual-motor performance. Such a skilled performance exhibits, according to Fitts (1962), three basic characteristics: (1) spatial-temporal patterning, (2) continuous interaction of response processes with input and feedback processes, and (3) learning. Much research

has been directed toward the development of laws which explain the operation of each of these characteristics of skilled performance. This research has resulted in the accumulation of a substantial amount of information concerning the nature of skilled performance and the processes by which skills are acquired. The information is of obvious relevance to the problem of training aviation personnel and therefore will be treated in this review. However, for several reasons these research results must be treated as information having implications for the problems in aviation but not necessarily being directly applicable. First, these studies were conducted in the laboratory environment. Although this allows the use of excellent experimental control during an investigation, it provides an altogether different context for the performance under investigation and may result in findings different from those obtained from research conducted in the air. Second, the normal stresses of aviation are absent, and it is well known that such stresses produce both quantitative and qualitative differences in performance. Third, laboratory tasks generally are less complex, both in terms of basic task dimensions and in terms of competing demands, than those found in aviation.

The above reasons dictate discretion in using the results of laboratory investigations in formulating general rules for aviation training. However, with these cautions in mind, this body of information may be used as an excellent source from which to generate hypotheses as to how aviation training might be improved and as to the nature of the research required to bridge the gap between laboratory findings and operational problems.

Many activities of a pilot can be classified as perceptual-motor. The manipulation of a panel of switches, the tuning of a radar set, the use of navigation equipment--all are subsumed under this class of performance. However, this section will be concerned only with continuous control activities, as exemplified by control of the aircraft in response to external visual cues or to visual information presented by a radar display. This is the perceptual-motor performance of most concern from a training point of view. In addition, the wealth of research concerned with human tracking behavior bears directly on this problem of aircraft control.

Placing the focus of attention on tracking behavior should not be considered restrictive in terms of understanding the human learning process. Fitts (1964) notes that "The theoretical framework within which skilled performance is now being viewed by most students of this topic is such that sharp distinctions between verbal and motor processes,

or between cognitive and motor processes, serve no useful purpose." Adams (1961) states that tracking studies, although frequently categorized as "motor skills research," should not be considered as separate from the so-called "higher mental processes." Senders (1959a) states, in fact, that "tracking behavior is such a universal aspect of biological functioning that it is really the underlying principle of all behavioral research."

A tracking task, in its simplest form, consists of a target, or error signal, typically presented through a visual display. The subject operates a control system through which he attempts to reduce the error signal to a zero value. The output of his control actions is indicated by a change in the display and, hopefully, by a reduction in the error signal. Controlling an aircraft is but one example of a continuous control task. Steering an automobile, riding a bicycle, and even placing a pen in a penholder are additional examples. One can see that the nature of the error signal and the characteristics of the controls can vary from very simple to quite complex. In fact, this variation in the complexity of the task has been cited by many as one reason for difficulties in systematizing much of the results of tracking research.

Research Directions Concerning Perceptual-Motor Capacities

Research concerning tracking behavior falls within several areas. Each area is concerned with a different aspect of the overall problem. Thus, each affords a somewhat unique contribution toward the development of a technology for aviation training. A brief description of these areas is presented below to provide a perspective within which to view the various contributions. More comprehensive reviews of the literature concerning the human as a manual-control element are provided by Summers and Ziedman (1964), Adams (1961), and Senders (1959a).

Frequency Analysis Techniques: The search for the human transfer function has attempted to develop a model which will describe mathematically the role of the human operator in a complex tracking situation. McRuer, Graham, Krendel, and Reisener (1965) state that "The description of human pilot dynamic characteristics in mathematical terms compatible with flight control engineering practice is an essential prerequisite to the analytical treatment of manual vehicular control systems." This approach, using the methodologies of servoengineering, compares the amplitude of an input frequency to the display with the amplitude of the same frequency and its harmonics as an output from the operator. Phrasing it somewhat differently, McRuer and Krendel (1957) say that if the characteristics of the human operator for a given overall task are

assumed to be capable of quasi-linear description, the operator mathematical model will consist of a linear transfer function plus an additional quantity inserted as an input into the system by the operator. The objective has been to develop a model of sufficient power that the additional quantity inserted by the operator, or the "remnant," will be a minimum. Excellent statements detailing the theory of this approach are provided by Walston and Warren (1954), McRuer and Krendel (1957), Licklider (1960), and Elkind and Green (1961).

Through use of frequency analysis techniques, several investigators have been successful in developing computer programs, termed model or analog pilots, which perform a tracking task in a manner identical to that of a human (Adams & Bergeron, 1964; Diamantides & Cacioppo, 1957; Goodyear Aircraft Company, 1952). Success of these programs is evidenced by the fact that the model can replace the pilot in the control loop without the pilot's knowledge of the change. To the extent that the model behaves like the human through a variety of system configurations, it affords a means for rapid evaluation of changes in control and display parameters.

Senders (1959b), using the results of a number of investigations, developed a set of optimum dynamics which might be used for the simulation of a one-dimensional aircraft tracking task. Frost (1962) conducted an experiment to assess the validity of the "optimum" control dynamics reported by Senders. For a single-axis tracking task, no difference was found between results using optimum dynamics and those in which subjects used a simple rate control. However, with a two-axis task in which the subjects were more heavily task loaded, the optimum dynamics produced reliably better performance than the rate control. Frost states that "use of the operator model rather than human subjects allows the human engineer to work directly with the structures and the flight control engineers to deal with stability and control early in the design phase of a system rather than wait until a design is frozen and only the cockpit displays are amenable to change." In a sense, then, this technique bears on the training problem in that it may be used as a means of eliminating at an early time control dynamics problems which might otherwise be solved only through additional training during the period of operational system use.

Certain objections have been raised against frequency analysis techniques as a means of providing a comprehensive understanding of human tracking behavior. Adams and Webber (1963) state that this mathematical theory "has never been able to accommodate the most commonplace psychological variables and human nonlinearities, and this is hardly surprising because the theory was devised independently of

behavioral data to describe machines. The human operator who attends, anticipates, learns, forgets, and fatigues can be considered a linear servomechanism only by weak analogy, although certainly the interest in servotheory has represented a commendable concern for mathematical description oi tracking."

It would seem, however, that the criticisms expressed above do not take fully into account the somewhat limited objectives of the exponents of frequency analysis techniques. Thus, Eakin and Campbell (1957) state that an important consequence of this approach would be an "increased understanding of the requirements of the aircraft and flight control servocombination that is best utilized by the pilot. Subjective evaluations often supply verbalized opinions of various system characteristics which aid the understanding of this problem. However, accurate measurements of the operation of the pilot in combination with aircraft systems present a much clearer quantitative picture of the interaction of the pilot with these systems. Such a knowlege should increase the ease with which aircraft and control system characteristics could be specified for a given weapon mission."

It is clear that investigators such as Eakin and Campbell feel that providing a statement of the transfer function of even an "idealized" pilot would be very useful for system design purposes. The extent to which the ideal pilot does not prove ideal during actual system operation would represent a source of internal system error to be studied and minimized at a later time, possibly through use of appropriate selection and training techniques.

Research Concerning Task Variables: Much of the research within engineering psychology is concerned with task variables. Birmingham and Taylor (1954) indicate a prime objective of engineering psychology is to understand the man-machine relationship sufficiently well that the higher order control requirements can be handled by machine components and the "transfer function required of the man is, mathematically, always as simple as possible and, wherever practical, no more complex than that of the simple amplifier." This philosophy has been implicit in many investigations of man-machine performance. Adams (1961) lists as representative samples of such task-oriented studies those concerned with control loadings, input signal characteristics, the magnitude of lag between control movement and system output, the effects of visual noise, mathematical transformation of the output signal, and compensatory versus pursuit tracking. Research by Chernikoff, Duey, and Taylor (1959) concerning the effects upon tracking performance of differing control dynamics in each coordinate of a two-dimensional task represents an excellent example of task-oriented research.

Task-oriented research has been productive and has been particularly successful in pointing out the dependence of operator performance upon specific characteristics of the system being controlled (Taylor & Birmingham, 1959). It also has produced system design philosophies, such as "quickening," which alter operator displays so as to provide him only the simplest information required to guide his responses. The resulting improvement in tracking performance is substantial.

Research Concerning Procedural Variables: A concern with variables that directly affect the proficiency of the operator represents the most traditional approach for the psychologist. Thus, tracking performance simply becomes one class of activity, among a number of classes, which serves as a dependent variable in the study of behavior. The influence of performance upon such variables as training schedules, motivation, fatigue, and physical and psychological stress represents the point of emphasis in such research. Here, as Adams (1961) points out, there may be little interest in the study of tracking for its own sake. He cites as examples of such research, studies on the rotary pursuit test with interest in fatiguelike effects or, more exactly, the implications of Hull's expressions of reactive and conditioned inhibition for behavior.

Any investigator who uses a tracking task simply as a means of eliciting a class of behavior upon which the effects of other variables may be studied should be aware of some severe and well-hidden pitfalls. Taylor and Birmingham (1959) illustrate results which can occur through the neglect of task variables and their influence upon measures of performance. They describe an analog computer demonstration in which the performance of three different "man"-machine systems were compared, using an amplifier in the place of the human subject in each instance. The only difference in the three systems was in the placement of the amplifier. The resulting system performance curves, shown in Figure 2, were all very different. The authors note that had these results come from a psychological laboratory, it might have been concluded that the subjects' performance improved most with the velocity system, next most with the position control device, and least with quickening. All conclusions would be incorrect inasmuch as the differing results are a function simply of a different system structure. The authors state that this shows clearly that the behavior of the system element cannot always be inferred directly from the performance of a system of which the element is a part. It would appear that the biggest pitfall here would be one of comparing, for example, two training regimes which used different tracking tasks in order to elicit the behavior.

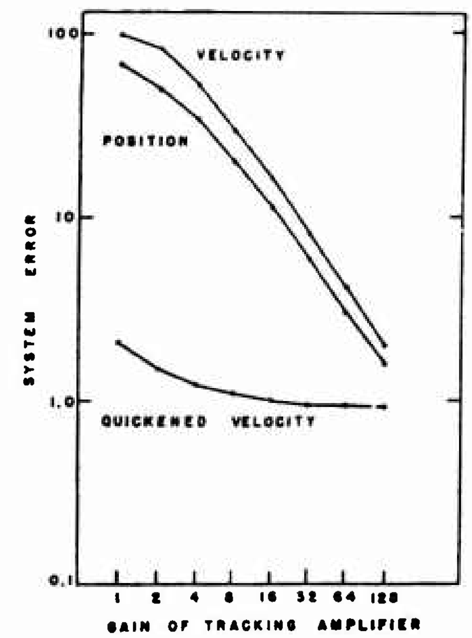


Figure 2. Performance of the Three Systems as a Function of the Gain of the "Man" (Logarithmic Ordinate). (From Taylor & Birmingham, 1959.)

<u>The Acquisition of Perceptual-Motor Skills</u>

It is generally agreed that the learning of a complex skill, such as the control of an aircraft, is a continuous rather than a discontinuous process. However, it may be advantageous to break this process into a limited number of arbitrary phases in order to compare the general character of the process during each phase. If differences are found to exist in, for example, the type of information required in different phases, the implications for the structure of efficient training programs are considerable.

The following discussion of the phases of acquisition of a complex perceptual-motor skill draws heavily on reviews of this process by Fitts (1962, 1964). Again, Fitts points out that the phase classification is an artificial one constructed merely for convenience of analysis.

<u>Early (Cognitive) Phase</u>: The early phase of skill learning, which may be very short for simple tasks, draws heavily on cognitive processes. The trainee attempts to "intellectualize" the task by structuring what is expected of him and the rules to be used during the learning process. He attempts to analyze the task and to verbalize about what is

being learned. The trainee seems to be developing a cognitive structure within which he can go about the actual practice of the task at hand.

The next major problem which arises for the trainee during the early phase is that of response integration. This is particularly important when a new task requires the simultaneous use of two previously differentiated sets of responses. This would be the case in the learning of a complex task involving both hands or hands and feet. The control of an aircraft obviously presents problems in response integration.

Intermediate (Fixation) Phase: This is the phase in which correct patterns of behavior are fixated by continuous practice until the probability of inappropriate response patterns or errors is reduced nearly to zero. The fixation phase, in a truly complex task, may last for weeks or months. Fitts estimates that in the case of the aircraft pilot, this phase would extend roughly from before initial solo through the time at which a private license is granted, and perhaps to the first hundred hours or so of flying.

During the intermediate phase, reliance on intellectualizing concerning the task drops rapidly until, it can be presumed, this verbal mediational process disappears entirely. Problems of response integration also diminish or disappear early in this phase.

Late (Autonomous) Phase: This phase, for which little experimental data are available, represents a very gradual, but continuing, improvement in proficiency. In addition, performance becomes more resistant to stress and to interference from other activities that may be performed concurrently. Fitts refers to neurological evidence indicating less and less involvement of cortical associative areas as learning continues in the case of simple conditioned-response learning, thus supporting the idea that this stage of autonomous behavior is based on a shift from reliance on visual to reliance on proprioceptive feedback, a shift of control to lower brain centers, and similar changes.

Peak performance in a complex perceptual-motor skill may not be achieved short of years of intensive, almost daily, practice. The fact that performance ever levels off appears, according to Fitts, to be due as much to the effects of physiological aging and/or loss of motivation as to the reaching of a true learning asymptote or limit in capacity for further improvement.

Characteristics of Perceptual-Motor Learning

There are certain characteristics of perceptual-motor learning which seem to hold regardless of the specific activity of skill involved. The following are the more important of these:

Continuous Nature of Process: The concepts of plateaus during the period of learning and terminal asymptotes at the conclusion of learning are now considered to be of dubious validity. Most curves showing the acquisition of a complex skill, while exhibiting considerable intertrial variability, maintain a rather regular progression to a point which might be taken as a terminal asymptote. Fitts (1962, 1964) cites three reasons why this apparent end to learning is not genuine. First, the nature of the task itself may place an artificial limit on learning, as in the case of typing, in which the mechanical lag of key operation places an upper limit on speed. Second, the criterion employed in measuring learning may not be sufficiently sensitive to reflect the small changes occurring after extensive practice. This would be the case for a time-on-target measure, with tolerances adjusted for the initial phases of learning. Third, there is ample evidence from such fields as music and professional sports that the learning process for highly skilled activities may continue for five to ten years or more.

Perceptual Versus Motor Aspects: In general, the performance of skilled perceptual-motor activities is considered to be more perceptual than motor. Fitts (1964) states that minimizing the role of motor behavior per se removes the principal basis for the commonly made distinction between verbal and motor processes. Instead, emphasis is placed on the intrinsic coherence of stimulus and response sequences and the cognitive or higher level processes that govern behavior sequence. This brings into importance such factors as timing, the interrelationships of speed, accuracy, and uncertainty, and the limitations imposed by capacities for discrimination and memory.

The conclusions of Fitts are supported by research of Parker and Fleishman (1960). In a factor-analysis investigation of the complex tracking task, it was found that predictor tests weighted heavily toward the motor aspects of performance accounted for less than 25 percent of criterion task variance. It was concluded that purely motor abilities do not determine individual differences in advanced tracking proficiency. These differences, which are substantial, appear to be related to perceptual aspects such as the ability to discriminate target motion and to predict target position at a future time.

Changing Ability Structure: Research by Fleishman and Hempel (1955) presents evidence for systematic and progressive changes in the basic abilities underlying proficiency in a perceptual-motor task with continued practice on the task. Figure 3 illustrates this graphically. This figure, based on results using the Discrimination Reaction Time Task, shows the more cognitive abilities to be of most importance initially, while those abilities having motor characteristics come into importance during later practice. Results such as these indicate that, while the acquisition of a complex skill may be a continuous and regular process, there are subtle quantitative and qualitative changes in the pattern of abilities determining proficiency at each stage of practice.

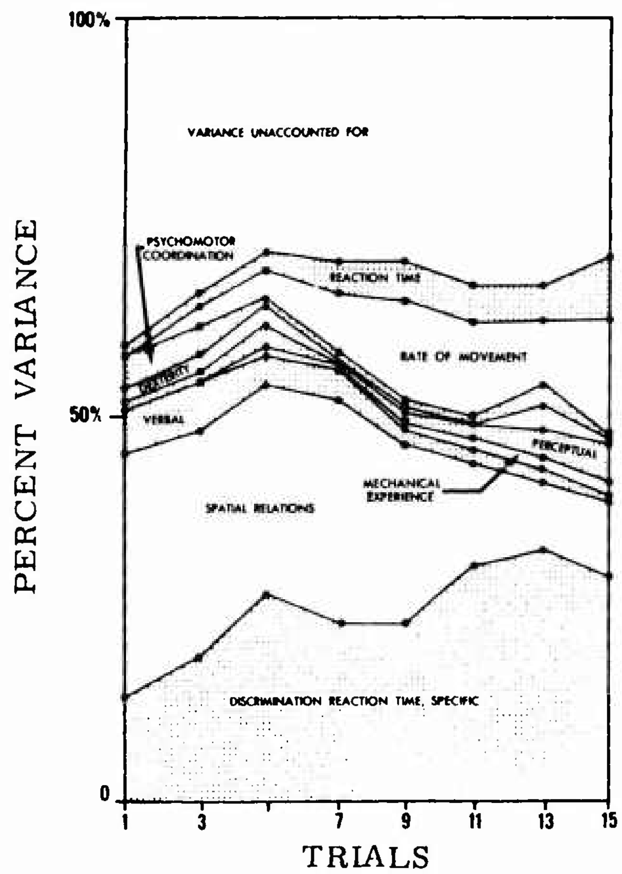


Figure 3. Percentage of Variance Represented by Loadings on Each Factor at Different Stages of Practice on the Discrimination Reaction Time Task. (Percentage of Variance is Represented by the Size of the Shaded Areas for Each Factor.) (After Fleishman & Hempel, 1955.)

Importance of Feedback: Adams (1964) notes that two classes of stimuli seem to be a necessary distinction for feedback research in motor skills. One class is reinforcing stimuli and is normally referred to as "knowledge of results." Most of the research accomplished on the efficiency of feedback information falls within this area.

The second class includes stimuli which through learning, come to provide the moment-to-moment regulation of behavior. These regulatory stimuli change continuously as a function of the continuous responses in the motor sequence and inform the individual about the rate, acceleration, amount, extent, and direction of movements. Although the operation of these internal regulatory stimuli is difficult, if not impossible, to observe, the elaborate motor sequences in a skilled perceptual-motor activity must be dependent upon the operation of some such guiding forces. In the paradigm of Adams, reinforcing stimuli operate to improve performance, whereas regulatory stimuli shape the learning of skilled acts.

It is obvious that without feedback of some kind, no skilled act can be accomplished. However, while the importance of feedback information is clear-cut, results of research attempting to manipulate feedback to improve performance or increase learning bring forth a number of contradictory conclusions. Karlin (1960) devised a task in which feedback information could be presented by means of visual, auditory, kinesthetic, and verbal cues. He found that no single modality resulted in superior learning or retention in comparison with any other. In addition, feedbacks which facilitated learning, in some instances impaired retention. For example, there was a tendency for visual feedback to be better for learning and poorer for retention. The reverse was true for verbal feedback. Inasmuch as the research of Karlin used a handwheel cranking task, a considerably different activity than that of concern in the training of pilots, the implication of his findings for aviation is questionable.

Most of the research within this area has been concerned with possible benefits from augmented feedback. Augmented feedback refers to providing knowledge of results through an additional feedback cue not inherent in the task itself. Smode (1958) presented additional feedback information in a form which was not sufficiently specific as to the direction and magnitude of error to permit subjects to use it as a cue to the making of specific responses. The increase which was noted in the performance of the tracking task was attributed to an increase in motivation rather than to added guidance.

Kinkade (1959), again using a tracking task, found augmented feedback to have a generally beneficial effect on perceptual-motor performance. However, the results of this study indicate that augmented feedback procedures must be used with some care. Kinkade found that (1) the relative change in amount of augmented feedback during training is more important than the absolute amount, (2) the optimum relative amount of augmented feedback is a function of the skill level of the individual, and (3) any change in amount of augmented feedback had a deleterious effect on tracking performance.

Garfinkle, Smith, Lyman, and Groth (1963) provided augmented feedback by magnifying the visual error by a factor of five and by providing proprioceptive cues related to both azimuth and elevation during a tracking task. Each technique of augmented feedback resulted in improved tracking performance. An interpretation of these findings must be tempered by a consideration of the results of the research of Buckhout, Naylor, and Briggs (1963). These investigators used a complex task which included both a discrete procedural activity and a continuous three-dimensional rate-control tracking task. Augmented feedback was provided for the tracking task through use of an auditory signal regulated by the amount of tracking error. No differences were found during either training or during a retention test after thirty days as a result of the use of augmented feedback.

Adams (1964) attempts to shed light on the contradictory findings concerning the effectiveness of auditory feedback through a consideration of the quality of the feedback provided by the task itself. He concludes that where the normal task feedback is clear-cut and unambiguous, augmented feedback is of little value. However, when task feedback is fuzzy or equivocal, augmented feedback provides a basis for increasing the precision of performance and for enhancing learning.

One final study is worthy of note concerning the effectiveness of feedback. Briggs and Wiener (1966) were concerned with the requirement for fidelity of proprioceptive feedback, or "control feel," during training. They were bothered by earlier results indicating that training task fidelity, in terms of control device characteristics (control loading), is not a necessary condition for adequate transfer performance in tracking-type tasks. Results of Briggs and Wiener indicate that high fidelity of proprioceptive feedback during training is important when the transfer task requires a relatively high level of time sharing, i. e., a complex task, but that with a relatively low level of such a requirement, it is not necessary to employ high fidelity in control loading. In view of the concern of the

present report for aviation training, the following conclusions of Briggs and Wiener should be noted:

1. A simulator to be employed for training in rudimentary flight control need not utilize a high fidelity of control loading in that (a) the time-sharing requirements are relatively low and (b) the level of skill attainable in such a device probably would not require that the subject employ proprioceptive cues to a significant extent.

2. However, control-loading fidelity becomes very important in simulators which (a) are used to train for skills requiring time sharing among a variety of displays and control devices, (b) are employed to provide extensive training, and/or (c) are utilized to maintain high levels of proficiency.

Training of Perceptual-Motor Skills

The following sections, which by no means review the entirety of the research information related to the training of motor skills, cite studies which are representative of the literature and list topics which are of genuine concern in the training of aviation personnel.

Training in a "Generalized" Tracking Skill: A series of studies supported by the U. S. Naval Training Device Center has been concerned with determining the existence and nature of a general tracking skill. Part of the impetus for these studies can be attributed to the traditional belief that the closer a training device resembles operational equipment, the higher will be the degree of transfer. In addition, it has been assumed that skills are specific to particular tasks and that trainers such as operational flight trainers must be quite specific to the aircraft of concern. This reasoning leads to the development of very expensive training equipment. If a general tracking trainer could be developed, the multiplicity of training devices could be reduced and a genuine economy effected.

In laboratory investigations, Kelley, Bowen, Ely, and Andreassi (1960) found that general tracking training produces significant increases in skill in the control of systems having widely different dynamic characteristics. These authors also found that transfer between different vehicular systems is high if instruments and controls are similar. Very stable high-order tracking systems seem, to the operator, to respond like low-order systems and provide an effective device for general tracking training.

Bowen, Hale, and Kelley (1962) attempted a field validation of the effectiveness of a general tracking trainer (the General Vehicular Research Tool). Using Army aviation trainees as subjects, it was found that on a final check ride, using a special evaluation form for measuring flight proficiency, the group given general tracking training was nine percent better on instrument flight and eleven percent better on contact flight. The three other criterion measures (regular school rating forms, time to solo, and Link trainer performance) all yielded negative results. The authors conclude that the principles embodied in a general tracking trainer appear to be sound. Practice of this sort does enhance performance in a specific vehicle. However, the general skill developed is difficult to isolate fully from the specifics of a given task. When these specifics conflict (in the training and operational situations), the gains from developing the general skill are at least partially obscured. Therefore, a given application may require special provisions for display and control features in the general tracking trainer.

Use of Part-Task Training: Much research has been concerned with the relative efficiency of training on the entire task or on part of the task (see also p. 163). Briggs and Naylor (1962) investigated the efficiency of several training methods for a complex tracking task. It was concluded that the best training procedure was some form of progressive-part training in which the part training became more complex as the subject became more proficient. Pure part and simplified-whole were not as efficient. It was reasoned that the progressive-part method utilizes a training task of high similarity to the transfer task and also provides an opportunity to develop efficient time-sharing behavior during the period of training.

Two points should be noted with respect to the above research. First, the progressive-part method, while superior to other methods of part-task training, was statistically equivalent to whole training, i. e., that involving the transfer task. Second, the initial transfer performance of the group trained by the pure part method was very poor and probably would have been judged "dangerous" had these training and transfer tasks occurred with real-life vehicular systems.

Effect of Environmental Noise During Training: Since most training tends to reduce environmental noise on the premise that learning is facilitated with clear-cut cues and in the absence of distracting influences, it is of interest to examine research information related to this premise. Briggs, Fitts, and Bahrick (1956) found that performance at all stages of practice with a complex tracking task was markedly degraded by the pres-

ence of visual noise on the steering-error dot. However, at the end of training, when groups were transferred to a mixed-noise condition, no significant difference was found in the performance of groups trained under noise versus no-noise conditions. This indicates that in spite of the marked performance differences, learning had progressed at approximately the same rate for the different training conditions.

Buckhout, Naylor, and Briggs (1963) obtained results in agreement with those cited above in which subjects trained under conditions of visual noise performed poorer than their noisefree counterparts during training. However, after a 30-day period without practice, subjects trained with visual noise were found to be superior when performing on a noisy display. Thus, although training performance is better without visual noise, it may be that training with noise is advised if individuals must perform in a noisy environment following long intervals without rehearsal.

Research Issues:

A considerable amount of research has been conducted in the general field of motor skills. The number of items listed in the Motor Skills Bibliography compiled by Ammons and Ammons for Perceptual and Motor Skills Journal attests to this. However, in examining this area from the point of view of developing a training technology for aviation, there are several major research issues which arise. The following, while by no means representing new insights, appear to be worthy of statement:

1. Field Evaluation Studies: Many laws concerning the acquisition of and the basic nature of skilled perceptual-motor performance developed in a laboratory using a tracking task as a criterion measure, appear to be meaningful for the training of pilots. A number of examples of such laws appear in this review. However, considerable caution must be used in extrapolating these research findings to the field training situation. Webb (1956) cites two studies in each of which one group was trained for a simple task, making the best use of laws of learning as developed in psychological laboratories. The program considered such aspects as distributed practice, knowledge of results, reminiscence, generalization, transfer, and the like. The second group was allowed to master the task in a free learning situation. Results were quite disappointing. In each instance, the free learning group proved to be superior. Webb states that "Through use of mnemonic devices, and who knows what besides, the free learning group actually learned more rapidly than the structured learning group." Webb, while conceding that there is truth in laboratory

findings, states that the translation of this information to the learning situation has been one of the most ineffective activities of the psychologist to date.

There is a need for a concerted effort to develop field studies using as a point of departure available laboratory information. These studies should not be merely attempts to replicate laboratory situations in the field. This would be rather futile. As Carstater (1956) points out, a frequent fault in the planning of training experiments is the failure to take the whole body of theory as a basis for an experimental design. The planning alone of an appropriate field validation program would be both lengthy and challenging.

2. Development of Appropriate Criterion Measures: It has been said in many different ways, but it still remains true: There is a real need for an effective criterion of good piloting performance. There is no effective and accurate way for measuring the proficiency with which a pilot plies his trade. Instructor's ratings, time-to-solo, and examination grades apparently are adequate measures for deciding whether or not an individual should continue in flight training. As shown in the research by Bowen et al. (1962), they are not adequate as criterion measures against which to assess the effectiveness of subtle variations in training regimes which might produce only minimal improvement in the efficiency of training. However, genuine progress in the training of aviation personnel may come only through a series of such minimal improvements.

Work on the development of a model or analog pilot appears to offer promise toward the development of a criterion measure for flight control skills. Once an analog pilot is developed, and these analogs are specific to individual pilots, characteristics of the transfer function for good and bad pilots (assuming the two ends of the piloting continuum can be obtained with some degree of accuracy) can be examined for any consistent differences. These differences, in turn, might suggest ways of instrumenting an aircraft to obtain objective and valid inflight measures of flight proficiency.

3. Generalized Tracking Skill: The quest for a general tracking skill appears to have been a commendable one and should be continued. The ultimate payoff would be considerable if training for a continuous control skill for a number of specific aircraft could be given in a single

general-skill trainer. However, the findings in this program, to date, indicate that more information must be known about possible interactions between the general trainer and the specific task before such a trainer can be used fruitfully. For instance, there apparently is some level of congruity between the controls and displays of the two devices which must be achieved. Further work toward the validation of the concept of a general tracking skill appears worthwhile.

RATIONAL ANALYSES

Rational analyses have been most artfully employed in defining the pilot's job in the operational context. The most prominent have been the task analysis techniques[4] which often use to advantage the opinions and experience of seasoned pilots. Task analysis is a method for describing the behavioral composition of tasks in a job situation (in psychologically meaningful terms, i.e., in terms of stimulus and response and the variables that intervene between these). It is one technique for determining systematically the behavioral requirements and skill dimensions involved in pilot performance. The basis for this is a listing of all the tasks the pilot performs for a defined aircraft/mission profile. This is followed by individual task descriptions, in operational terms, of the activities

---

[4]Task analysis is a tool for analyzing certain aspects of man-machine systems to achieve defined purposes. In the design of training systems, task analyses have been conducted for purposes of: specifying training requirements, training equipment design and use, and developing instructional programs (in pilot training prog ams, for example, the pilot job requirements are also defined). Task analysis is also conducted for human engineering and systems design purposes as well as for selection procedures and criterion test development. Depending on the specific purpose of the analysis, various methods of presenting this kind of information are employed, ranging from tabular to symbolic representation and involving various levels of complexity, beginning with simple lists of desirable traits in an individual to highly specific time and motion analyses. Most often, these are based on observation of an individual's performance. Some well-known formats used in military systems include the Task-Equipment Analysis (TEA), Operational Sequence Diagram (OSD), and Information, Decision, Action (IDA) charts. A great many human factors scientists and engineers have contributed to the development of these formats, and much has been written on task analysis (see B.J. Smith, 1965; Wright Air Development Division, 1960).

performed, usually represented by a time and event linkage of task and task elements (both sequential and coordinative activities) per flight segment, the equipment involved (displays and controls), decisions and actions, response feedback, frequency and duration of performance, criteria which define adequate performance, and conditions under which the activity is performed. From this, the behavioral requirements of the tasks in terms of knowledges, skills, and concepts (i.e., task analysis) needed to meet the performance requirements are identified.

Considerable differences exist in the literature as to the definition of task descriptions and task analyses (B.J. Smith, 1965; Chenzoff, 1964). Conventionally, the task description refers to a detailed listing of the operations the pilot and aircrew perform in a flight mission. Task analysis refers to the ordering of the task descriptions into behavioral terms, i.e., describing the behaviors required in the detail needed for the definition of skill and knowledge categories which have relevance to training.

Task analyses of some sort have been conducted for many aircraft/mission profile combinations to determine the activities involved in a pilot's job.[5] This derived information provides the means for achieving various aspects of training system design, e.g., anticipating training requirements, preparing training recommendations, and specifying training equipment. Task analyses also represent the most useful way available today for defining the pilot's job in the mission context.

Experience with task analyses has made clear the fact that their content is not generalizable across training situations. A task analysis is tailored to the needs of a specific situation--hence the lack of universality in format. However, the basic procedures, information categories

---

[5] Task analyses of varying levels of emphasis and detail have been prepared for all recently developed major aviation weapon systems, and military specifications have been published for providing guidance as to their form, content, and time of preparation in the development cycle (see HIAPSD 80-3). A recent example of a task analysis for establishing crew procedures and division of tasks between the crew has been published for the F-111B aircraft by Grumman Aircraft Engineering Corp. (1965).

sought, etc., tend to be similar in training system development. Figure 4 shows a task analysis format that has been used by the Air Force (Snyder, 1960).

As mentioned earlier, the distinguishing feature between the task description and the task analysis is the added operation (which defines task analysis) of describing human behaviors required by the job in psychologically meaningful terms. Unfortunately, the process of determining the behavioral requirements is complex and difficult. The available procedures are inexact and cannot be applied in a standard way by all people. Specifying the knowledge and skill requirements imposed on the pilot for a given aircraft/mission profile is a highly subjective process and is based in large part on the experience of the analyst. What usually is accomplished under the guise of task analysis is a detailed description of pilot activities, in operational terms, in a time and event sequence analysis. Thus, techniques which do not belabor the behavioral orientation are heavily used to depict human activities in a system in operational terms. The format used in defining the pilot(s) activities usually indicates: the purpose of the activities performed, the equipment involved, the conditions under which the activities are performed and the relationship of a given activity to other tasks at the same or related positions in the system. Commonly, the activities are presented in their normal or logical sequence of occurrence. Time information may be provided by a time scale or by statements indicating the percent of total time spent in performing an activity. For a flight mission, the sequence of description begins with flight segments, each of which is made up of tasks. Within each task are the task activities or elements which define the performance. In present-day aircraft this type of analysis becomes extremely complex. An excellent example of this complexity is afforded by the task analysis accomplished for the Boeing 727 jet commercial airliner.[6] The number of tasks and task elements (a task is a limited and orderly grouping of purposeful activities; elements are the activities which make up the task) performed by the crew for each segment of a normal commercial flight is shown below. While this analysis was done for the purpose of developing a training program for 727 crews, the pilot job requirements were explicitly set forth.

---

[6] Analysis conducted by United Airlines, Denver, Colorado.

TASKS (GROSS TASK) (SUB-TASKS)	VEHICLE EQUIPMENT	SUPPORT EQUIPMENT AND PERSONNEL	DISPLAYS	DIAGNOSIS AND DECISION	ACTION	FEEDBACK	INCIDENCE	TIME	REMARKS
Take-Off (Example of Gross Task)	Throttle, Stick, Rudder Pedals	Control Tower	RPM Instrs, Gyro Instrs. T.P.T.	Recognizes that steps preliminary to take-off are satisfactory	Pushes Throttle Forward	Increased RPM: Appearance of Runway; Sound of Engine	Once Per Sortie	30 Sec	
Radio Position Report (Example of Sub-Task)	ARC ____	Communica-tion Station	Pilot Light or Switch Position, Clock	Recognizes Time for Report	Depresses Mike Button	Clicks In Earphones; Verbal Responses from Ground	Every ____ Minutes	10 Sec	
Identification of task a. Include alternate tasks b. List tasks in operational sequence	Equipment task is performed on.	Personnel performing task--number and type. Other personnel or equipment supplying information or services.	What operator sees, hears, feels, smells, etc. This must include all the critical elements, needed for decision, including controls which may also serve as displays.	(1) Mental and/or physical process before initiating action. (2) Consideration of alternates (other possible decisions).	What does operator do? Consider possible erroneous actions. Define the control(s) involved, their location, and tolerance.	What indications does pilot get of effect of his action?	How often is pilot called on to perform this activity?	How long does this action take? Consider overlap or interference with other actions and total time requirement.	Degree of confidence on task existence during operational procedure.

Figure 4. Air Force Task-Equipment Analysis Format (from Snyder, 1960).

Segment	Tasks to Perform	Task Elements to Perform
Preflight	164	625
Engine Start and Taxi	63	281
Takeoff	10	26
Climb	36	103
Cruise	35	172
Descent and Hold	36	99
Approach	92	125
Landing	26	76
Taxi and Shutdown	34	132

An opportunity to evaluate the adequacy of the task analysis method for deriving training requirements was afforded United Airlines Flight Training Center personnel.[7] Training requirements derived from task analysis data from Boeing 727 aircraft were compared with the B-727 transition program conventionally developed by UAL. The conventional program consisted of pilot-qualified professional instructors' defining the training requirements and organizing them into course outlines and lesson plans for classroom lecture, crew training with systems mockups, flight simulator training, and actual aircraft flights. The training experts had the benefit of reviewing all available flight test data and were trained by factory instructors and test pilots. Also, the developed program was given a dry run by UAL pilots and instructors for debugging prior to the first line crew classes. Results indicated that the task analysis did not generate any more complete or more thorough training requirements than were already in the transition program. The task analysis did not serve as a substitute for expert training judgment. However, the qualification was added that the personnel who developed the conventional program drew upon more than 15 years of experience with previous aircraft programs, whereas the complex task analysis study was the airlines' first experience with the method. With more experience with the method, it was felt that greater benefits would accrue. The task analysis exercise did focus on the question of what are the essential job knowledges rather than on how the airplane is put together, and this was viewed as a positive contribution that could conceivably improve future training programs particularly with new additions to the fleet.

---

[7] Weaver, F.L. The Boeing 727 Task Analysis and the Boeing 727 transition program--a comparative study. United Air Lines Flight Training Center, Denver, Colo. October 1964.

The task analysis difficulties in defining the job of the pilot are part of the more general issue of developing concepts and methods that are adequate for preparing useful and detailed training requirements. Present methods for achieving this are weak. What is needed is a taxonomy for ordering tasks in such a way that applicable principles of training can be specified as optimum for a task class.

A number of investigations have attempted to develop task classifications, but these have turned out to be less than adequate for purposes of training. The need for a taxonomical structure as a prelude to the specification of training procedures (including training equipment specification) is well understood. Task classifications currently available are best described as initial attempts. Although various schemes have been proposed for operator tasks, almost all embrace similar sets of behavioral components. (A discussion of the problems in task taxonomy is presented in Section IV of this report).

CORRELATIONAL STUDIES

Information pertinent to the definition of the pilot's job has also been provided by factor analysis studies of flying performance. The basis for this approach has come from a program of experimental laboratory studies of ability variables underlying perceptual motor tasks (see Fleishman, 1962). Essentially, the premise for training research is that the ability categories developed in the laboratory researches may be used in describing performances in the pilot training situation. Thus, intercorrelations among measures of flight control performances provide a basis for inferring the nature of the skill dimensions underlying the pilot's job in the perceptual-motor domain. "Factor analysis techniques applied to such intercorrelations yield a mathematically defined set of dimensions descriptive of skills common to the diversity of subtasks involved in the total task [of flying an aircraft]." (Zavala, et al., 1965)

Pilot Research

Two studies using this methodological approach bear directly on the pilot's job. Fleishman and Ornstein (1960) conducted a factor analysis of performance obtained on 24 different contact flight maneuvers performed by 63 Air Force student pilots in the T-6 aircraft. The measures were obtained from the Daily Progress Record Sheets (DPRS, developed by Smith, Flexman, & Houston, 1952) made out on each student by the instructor. A separate DPRS was completed for each maneuver, and each item within the maneuver was scored "correct" or "incorrect."

The sum of the incorrect items (both subjectively and objectively determined) was the maneuver error score used in the analysis. The factor analysis yielded seven factors. Attempts to interpret these in terms of common subtask operations in flying or common control movements were not fruitful. The factorially complex maneuvers (a composite of tasks) were difficult to define in behavioral terms. What best fitted the data were psychomotor ability factors derived from the ability model developed from experimental-correlational analyses of laboratory perceptual-motor tasks. Thus, the factors describing the common requirements of the aircraft maneuvers were identified in such gross terms as: Control Precision, Spatial Orientation, Rate Control, Multilimb Coordination, and Kinesthetic Discrimination.

Within the same framework, a recent study analyzed pilot performance in Army helicopter flying (Zavala et al., 1965). Factor analysis techniques were applied to the intercorrelations among maneuver proficiency scores obtained from flights of helicopter pilots at the completion of primary and basic stages of training. Data were obtained on 538 students in primary flight training and on 383 students in basic flight training. Proficiency data were taken from the Army Pilot Performance Description Records (PPDR). The PPDR, developed at Fort Rucker, Alabama (Greer, Smith, & Hatfield, 1962) provides for scoring of specific components that make up a flight maneuver, and for summary evalautions, which are overall ratings made by the instructor for each maneuver on a 4-point scale (above average to unsatisfactory).

Pilot performance was analyzed into independent component abilities. The common factors identified in the primary phase analysis were:

Takeoff
Autorotation and Forced Landing
Hovering
Traffic Pattern
Forced Landing from Hover
Landing

In the basic phase analysis, the factors were:

High Reconnaissance
Forced Landing from Hover
Slope Operation
Takeoff Preparation
Low Reconnaissance to Landing

As in the Fleishman-Ornstein study, similar difficulty was experienced in interpreting the factors in terms of general skill categories. These "molar" qualities were not readily translatable into behavioral categories, in part due to the complex sequence of tasks involved in each maneuver. To further refine the interpretation, the authors examined intercorrelations among items selected from over 200 individual task scores from the pilot proficiency data. The interpretable task factors identified for the primary phase were:

Airspeed
Pitch Application in Forced Landing from Hover
Line and End of Descent
RPM
Amount of Pitch
Airspeed Reduction and Rate of Descent
Rate of Closure
Power-Off Pitch Application
Downwind Airspeed

The interpretable task factors identified for the basic phase were:

Drifts
Low-Altitude RPM
Rate of Closure
Confined Area Spatial and Angular Judgments
Amount and Timing of Aft Cyclic Without Power
Power-Off Pitch Application
Airspeed and Airspeed Reduction/Increase
High-Altitude RPM
Observation Angle of Sight
Low-Reconnaissance Descent Angle

In addition, six factors common to the two analysis were identified. These are listed together with hypotheses formed by the authors about the basic abilities required for each task.

Common Factor	Hypothesized Ability Dimensions
Engine RPM	Division of attention Precise hand-wrist movements Coordination of throttle, cyclic, & pedal movements
Airspeed	Division of attention Coordination of arm movements with display cues
Line, end, and angle of descent	Integrated motor and spatial abilities
Rate of closure	Fine controlled arm movements coordinated with visual motion cues
Power-off pitch application	Precise arm movements; timing in judging distance and rate and integrating them
Drifts	Fine simultaneous arm and feet movements

It is difficult to determine what the data from the factor analysis studies offer for pilot training programs. Not much is available that bears usefully on the definition of components in the pilot's job. Actually, the available evidence has come from the study of ability dimensions in perceptual-motor performance employing experimental-correlational analyses of laboratory tasks. In these investigations, the criterion measure or the performance being studied has not been as complex as that found in the real world. Thus, the value of this type of data for improving pilot skill training is questionable, at least in the present form of development. A number of the derived ability factors leave one puzzled as to exactly what implications these may have for training sequences or how they can be translated into effective training regimes. For example, the factor of "multilimb coordination" is well understood by anyone in aviation as an ability requirement and one that underlies all manual control skills in the air, i.e., it develops in qualified trainees with practice, total training, and experience. Tieing this factor in with related events (e.g., sideslip component, cross-controlling, coordinated turns, etc.) is truly an intuitive undertaking. Assuming that these ability factors are descriptive of laboratory tasks, it appears that their power of description decreases exponentially as one moves to the complex

flight tasks of the real world. The question is, what is the instructional meaning of such ability factors? Perhaps one answer to this is the attempt to make specific use of ability requirements information in designing defined training sequences.

Use of Task-Specific Information

There is evidence that ability patterns postulated to underly proficiency at different stages of learning can be used to guide the student during successive training sequences. Parker and Fleishman (1961) investigated the extent to which the effectiveness of training for a complex tracking task might be increased through use of training procedures based on information as to ability factors which were important contributors to proficiency at different stages of practice. Thus, if it was known that some ability factor, such as multilimb coordination, or some component performance, such as rudder control, normally came into importance at a specific point in the training schedule, an attempt was made, through verbal instructions, to emphasize that feature at a slightly earlier period. The experimental training program was compared with a program in which subjects received no formal training other than the answering of questions, and a "common-sense" training program using standard pedagogical techniques. There was an indication of consistent superiority for the experimental procedures throughout the course of practice (time-on-target performance). Thus, pilot training might benefit from this approach, particularly on those identifiable task components where a trainee is weak, even if the components were isolated from the larger context in which the component is imbedded. Perhaps practice on common components may also generalize to a larger variety of flight maneuvers. This may be incorporated with the more general notion of stimulus support training (Skinner, 1958). Providing the trainee with ample stimulus cues and prompts during initial training increases the prospect that desired responses will be made with a minimum of errors. These extra stimulus cues and supports are removed as the trainee moves closer to the desired behavior until the operational situation is approximated. Similarly, the trainee may make better use of performance information given to him in terms of scores on task factors if these can be made explicit and differentiated from information on proficiency in a maneuver given in more general terms.

However, at various stages of training, presenting the trainee with instructions based on intuitive definitions of factors appears to be a tricky business and the reliability of this technique can be questioned. It is exceedingly difficult to obtain comparable measures of component abilities, and the timing of their relative importance for piloting tasks,

without information as to the time at which each ability comes into importance as proficiency is achieved. At present, there is little use to be made of this type of information in structuring a training program.

One may question the factor analysis method in still another sense as to its utility in deriving components of the pilot's job. In the studies cited on pilot factors, rather elegant statistics are applied to inflight performance scores that are somewhat lacking in reliability and are of unknown validity. From this, expert but intuitive judgments are made in factor identification. Would similar conclusions obtain from a range of scoring instruments possessing differing reliability and validity? More study is needed to answer this objection. Thus, the important research questions, so far as pilot training is concerned, center on two issues:

> How adequate for pilot training needs is the taxonomy of factors derived by factor analysis methods?
>
> How effectively can the taxonomy be used in pilot training situations?

Research Issues:

Knowledge concerning the definition of the components that make up the pilot's job and the behavioral requirements that underlie these components is in real need of extension. Current methods are not sufficient to define the job in the detail needed for training purposes in terms of relevance to all classes of weapon systems and mission requirements in Air Force operations. The published research defining the pilot's job has been sketchy and inadequate. In large part it has centered on continuous manual control tasks in the laboratory setting and on inflight control and positioning of the aircraft for a variety of school maneuvers. Little concern has been given to the many other task requirements in flying or to the simultaneous and integrated aspects of task performance.

Knowledge on how components of flying performance are learned is a prime topic for research. There is a need for parametric studies which will generate a series of learning curves for various flight skills, i.e., generalizable data on how rapidly the pilot acquires flying skills. Systematically developed information of this sort is of substantial value in determining specific levels of performance required during stages of training and also in determining the optimum flight curriculum.

Present evidence is sketchy concerning what it is that distinguishes one flight skill from another. Thus, a major part of this effort should center on a realistic and precise identification of critical, trainable job components and tasks. Initial support for this venture should come from the previously mentioned data derived from perceptual-motor analyses and pilot expertise on specific flight vehicle operations in mission environments, and from the flight components identified in the factor analysis studies.

Experimental study should continue to investigate the effects of using ability components or critical job component information in the early stages of flight training. This is based on the suggestion that the efficiency of initial training in a complex tracking task can be optimized by utilizing analytical information about ability requirements during the course of instruction, i. e., certain abilities are important at given points in the learning of skills, and training can be structured to emphasize these at appropriate times during training. For example, initial training in a light aircraft or in a simulator, emphasizing instruction in critical skill components at specified times, can be compared to the more conventional approach in terms of time required to reach the criterion for the phase of training, and also in terms of the effects of this training on transfer to more complex flight maneuvers.

SECTION III

RESEARCHES RELEVANT TO PILOT TRAINING

This section presents an appraisal of the considerable number of studies dealing with the acquisition, assessment, and retention of flying skills, and with behaviors pertinent to the pilot's job. Three major areas of flight operations are considered in which researches on a variety of aspects are reviewed and evaluated.

The first grouping deals with topics subsumed under skill acquisition. One portion is devoted to studies investigating the effects of sequencing of flight training on the efficiency of skill acquisition of student pilots. Another portion appraises the research on the critical and important components of operational flight tasks. The last part concerns the measurement of pilot performance.

The second grouping of studies centers on simulation and transfer of training. Two portions deal respectively with researches on the effectiveness of simulator training, and simulation requirements for training.

The third grouping of studies considers the maintenance of flying proficiency, specifically treating the retention of flying skills, and decrements in performance over time.

The specific research areas reviewed in this section are organized as follows:

<u>The Acquisition of Skill</u>

- Contact/Instrument Flight Training
- Light Plane Training
- Manipulation of Instructional Variables

<u>Training in Operational Components of the Pilot's Job</u>

- Low-Altitude Flight
- Pilot Workload
- Performance and Stress
- Crew Training
- Visual Aspects in Flying

Performance Measurement

Development of Objective Flight Checks
Quality Control
Assessing Aptitude for Military Flying
Scoring Capabilities in Simulators

Simulation and Transfer of Training

Effectiveness of Simulator Training
Simulation Requirements for Training

Motion Simulation
Extracockpit Visual Simulation
Part-Task Trainers
Fidelity of Simulation

Maintenance of Proficiency

Retention of Flying Skills
Performance Decrement Over Time

THE ACQUISITION OF SKILL

The implications of sequencing flight training on the efficiency of the skill acquisition process are the topic of this part of the report, and researches are evaluated which are germane to improving training in initial flying skills and to improving training in ground school. Three groups of studies have been selected for review. The first group of studies investigated the merits of the sequencing of contact and instrument flight instruction on flying training. The second group of studies is concerned with the effects of preprimary light plane training on the subsequent acquisition of flying skills in heavier military aircraft. Finally, studies are reviewed which consider the effects of manipulating course and instructional variables on initial acquisition of flying skills.

Contact/Instrument Flight Training

Sequencing contact and instrument training on initial skill acquisition has been the subject of a number of research reports. These studies have investigated the issue from two different aspects. The first considers the effects on flying performance of giving all of one kind of instruction (contact or instrument) to the trainee before giving any of the other. The second considers the effects of giving instrument training early in the flying training program by integrating it with contact training,

and also explores the simultaneous training of pilots in the use of cues from both visual sources. Each of these aspects of the problem is discussed below.

Sequence of Training: Flight training programs traditionally have followed a sequence in which the student is first required to demonstrate a degree of mastery on contact flight before being given any instrument instruction. Thereafter, instrument-only training is given for a block of time. This sequence has been accepted quite naturally over the years because of the manner in which aircraft have evolved. In the beginning, the pilot flew the aircraft by feel and was in constant contact with the ground. Instruments were meager, and nav-aids consisted largely of a compass and features of the terrain such as railroad tracks, rivers, etc. As instrument development increased, training with these equipments was usually "tacked on" after the pilot soloed. Experience with this traditional order of pilot training has indicated, however, that it is not maximally efficient nor even desirable for initial training. The reasons given are that it may (1) force the student pilot to develop habits which make it difficult to learn instrument flying techniques, (2) produce pilots who, although instrument qualified, lack confidence in instrument flying techniques, and (3) fail to provide emergency instrument training for the 30- to 40-hour pilot (see Jolley, 1958). Some aviation training experts (e.g., Eliasson, 1961) have accepted as fact that students with prior contact experience take longer to learn instrument flying than students with no previous flying experience. What evidence can be marshalled in support of this?

Ritchie and Michael (1955) performed an experiment to determine the effects of instructional sequence on the acquisition of contact and instrument flight techniques. Two groups of student pilots (n = 11 each) were trained to criterion on two flight maneuvers. Subjects learned to fly a Piper "Tri-Pacer" airplane on a straight and level course and to make level 180° turns. One group learned contact first and the other learned instruments first. Contact flying was easier to learn than instrument flying (fewer trials required), and contact and instruments had different transfer effects on each other. Flight by instruments was harder to learn when contact flight was learned first; but, instruments-first facilitated contact learning. Students who learned contact first required 22 percent more trials to acquire instrument techniques than the group learning instruments first. The group learning instruments first, however, required approximately 47 percent fewer trials to master contact flying. The authors state that the difference in the direction of transfer "would be expected to reduce consistently the total learning time for both tasks when the instrument task is learned first," but this finding was not conclusively established. It was concluded, however,

that the traditional order of presenting flight instruction allowed the student pilot to develop habits which made it unnecessarily difficult for him to learn instrument flying techniques.

Ritchie and Hanes (1964) replicated the Ritchie and Michael study to verify the existence of the positive (instruments to contact) and negative (contact to instruments) transfer of training[8] found in the earlier study. Two groups of university students (n = 13 each) were trained to fly the maneuvers to the same criterion used previously. As before, one group learned instruments first and the other learned contact first.[9] Again, contact flying was shown to be the easier task in terms of time to learn. Ritchie and Michael's finding of positive transfer from instruments to contact was verified. The group who learned instruments-first learned contact in 53 percent fewer trials than the contact-first group. The finding of negative transfer from contact to instruments, however, was not verified. In this study, learning contact first reduced the number of trials to learn instruments by 20 percent (i.e., there was a +20 percent transfer to be compared against the -22 percent found in the first study). On instruments, straight-and-level was the difficult task to learn and heading was the limit most often exceeded. In contact flight, turns were more difficult, with no one limit exceeded--in bank, airspeed, or heading--accounting for the majority of errors. The Ritchie and Michael data suggested that total learning time (contact plus instruments) could be reduced by presenting instrument flying lessons first in training. Ritchie and Hanes found, however, that task order had essentially no effect on total learning time (difference of only three trials).

These two studies demonstrated that, of the two flying techniques, contact can be acquired more quickly than instruments (i.e., is easier to learn). Instrument training given first in the flight sequence certainly does not "interfere" with the subsequent acquisition of contact skills.

---

[8]Transfer of training is here measured in terms of the number of trials required to reach the criterion on the transfer task. If more trials are required as a result of first-task experience, transfer is said to be negative; if less are required, transfer is positive. Perhaps a more accurate term would be positive or negative "savings."

[9]A third group who learned instruments in a Link Trainer prior to transferring to the aircraft was included in the replication. The performance of this group is not directly relevant to the present discussion. These results are cited in the discussion on Simulator Training, page 146.

The data also suggest, for the maneuvers and aircraft used, that contact learning time can be reduced significantly following instrument training. However, there is no evidence that total learning time can be reduced by reversing the traditional training sequence and presenting instruments first. The question of whether contact instruction given first in training makes instrument learning time longer for the pilot trained in this way than for a student with no previous flying experience is unresolved, since the trends observed in the first study were reversed in the second.

A key factor that must be considered in evaluating studies of this nature is the actual design of the flight instruments used. It is well known that many flight instruments are difficult to read and interpret. Ritchie and Hanes (1964, p. 28) make special note that 96 percent of the difference between contact and instrument flight conditions on the straight-and-level maneuver was due to inability to hold heading. They further note that the directional gyro used in the experiment was a standard World War II instrument that had been criticized as a display and replaced by the Air Force, Navy, and commercial airlines. The extent to which instrument design affects the learning of instrument flight techniques must, of course, be taken into account when evaluating the relative difficulty of learning contact versus instrument skills. Clarification is needed as to the extent to which instrument flying is more difficult than contact flying due to the artifact introduced by instrument design.

A further qualification was placed on these studies by the maneuvers used as criterion tasks. By no stretch of the imagination can one accept a simple turn and holding straight and level as representing the complex of instrument flying skills or, indeed, contact flying skills. Attempting to generalize from this limited information base is futile. What is needed to put the sequencing problem into operational perspective is a substantial program in the military context, utilizing current instrument displays.

Integrated Training: The suggestion is plausible that there is no apparent reason to defer instrument training until contact flying has been mastered. Teaching the pilot early to fly on instruments may well overcome later reluctance to rely on and use instruments when alternatives are not available. It may also improve the overall quality and efficiency of flight instruction. A number of studies subsumed under the general topic of integrated contact and instrument flight training have investigated the effects of early instrument training on subsequent flying proficiency. Essentially, two ways exist in which contact and instrument training may be integrated. The first refers to integration within the training program, where instrument training is given early in Primary (i. e., beginning phase of pilot training) and integrated with contact instruction. The second is integration within a lesson, where the trainee is taught to

use both sets of cues simultaneously. Both training concepts are discussed below.

A program of study performed by the Navy (Creelman, 1955b; 1955c; 1955d; Creelman, 1956) sought to determine the training effectiveness of a proposed modified flight training syllabus wherein students would receive instrument training distributed throughout the syllabus rather than concentrated at the end. Instrument training (16 basic instrument instruction flights) began at the acrobatic stage of the primary phase for an experimental group (n = 50). The regular course was given to the control group. Performance of the two groups was compared at several points during later training. Results of the comparisons at the selected points, reported in the above-cited documents, were:

1. On measures of acrobatic stage proficiency during Primary (A, B, and C Stage UBAA grades[10] and number of "downs"), no significant differences existed between groups. The conclusion was that the modified instrument syllabus did not interfere with overall flying ability as measured (Creelman, 1955b).

2. On 12 gross measures of noninstrument flight proficiency during Basic Training (e. g., Formation Grade, Basic Flight Grade, Cross-Country Navigation, Field Carrier Landing Practice, Carrier Qualification, etc.), the experimental group was significantly better on three measures (number of days from C1 to end of basic, solo hours, Primary Combat Grade), and differences were negligible on the other nine. The results of the comparisons made indicated that early instrument training had a facilitating effect on trainee proficiency, since the experimental group had fewer solo hours in which to practice but achieved the same level of proficiency (as measured by the Basic Flight Grade) as the control group (Creelman, 1955c). The conclusion offered was that the integrated instrument procedure could be introduced during acrobatic stage with continuing refresher hops during subsequent phases without interfering with non-instrument flight proficiency.

3. On measures of proficiency in the night primary and instrument stages (e.g., number of downs, above and below average marks, night primary grades, etc., and analysis of instructor opinions), the experimental group's performance was significantly poorer than the control

---

[10] UBAA grades are average grades derived from the sum of Unsatisfactory, Below average, Average, and Above average grades, divided by the total number of marks.

group's on night primary grades and number of downs. Overall instrument grades were essentially the same for both groups. Flight instructors felt that the integrated syllabus would work optimally if the first night instrument flights were given immediately after, rather than during, the aerobatic stage (Creelman, 1955d).

4. On measures of advanced instrument proficiency (e.g., Phase A Flight Grade, number of rechecks, etc.), there were no significant differences between groups. Thus, the author concluded that the integrated instrument syllabus had no effect on the available measures of advanced instrument proficiency (Creelman, 1956).

The University of Illinois, Institute of Aviation (1955), performed a study to determine the feasibility of incorporating both instrument and contact flight training within the private pilots' syllabus. The purpose of the experiment was to train, within the allotted 40 hours, private pilots who could meet contact flying requirements and who could also demonstrate positive control of the aircraft on instruments. The first five training periods (total of 3.2 hours) were spent on instruments. In the sixth period, contact flight was introduced and thereafter contact and instrument flying were interspersed where possible. Upon completion of the course, all 18 students passed the private pilot's test (contact) and demonstrated "appreciable" ability to fly on instruments. A similar study at West Virginia University (cited by Jolley, 1958), reported essentially the same results under similar conditions.

Ritchie and Hanes (1964) cite an unpublished Army Aviation study by Prophet, conducted in 1963, which found that student pilots given instrument training early in flight training passed 30-hour check rides with better scores (and fewer dropouts) than students in a control group. Differences in performance, however, tended to disappear after approximately 200 hours of flight experience. Interim data, in a report of progress by the Human Resources Research Office,[11] supported the general conclusion that integrated training produces students who exhibit slightly better performance during primary training, but the advantages disappear by the completion of advanced contact and instrument training.

In the integrated studies described in the foregoing, the student first learns to perform a maneuver(s) by use of either contact or

[11] Human Resources Research Office. Quarterly Progress Report, April-June 1964. U.S. Army Aviation Human Research Unit, Fort Rucker, Alabama, 30 June 1964.

instrument cues (but not both) and then relearns it by use of the other set of cues. This type of training has been called "block" training to distinguish it from "true" integrated training. True integrated training involves the simultaneous presentation to the student of both contact and instrument cues while he is learning a maneuver. During this type of training (i.e., "simultaneous cue" method), the instructor systematically demonstrates to the student how the instrument and contact cues relate to each other.

The efficacy of this method stems from the observation that the experienced pilot flies his aircraft in an integrated fashion by use of cues from either or both visual sources alternately, simultaneously, or combinatorially. Thus he controls the aircraft by means of a single integrated flying technique rather than by instrument or contact flying techniques (Poe, Jolley, & Prophet, 1960). This integrated training concept seeks to teach the pilot the "single technique" of aircraft attitude control rather than waiting for him to develop it on his own.

Apparently, no valid test of this training concept has yet been accomplished. An unpublished Air Force study (cited in Jolley, 1958), done at Graham AFB in 1956-1957, reports limited data regarding integrated instrument/contact training. Two primary pilot classes were trained employing integrated concepts. Differences between integrated training groups and control groups (flight checks) were small but in favor of the integrated groups. No definite conclusions could be reached because of uncontrolled variables operating in the experiment. Instructors, however, did note that learning to use two references at the same time (simultaneous cue method) imposed too much of a burden on the beginning student.

A feasibility study of an integrated training program aimed at producing instrument-qualified Army aviators was conducted in 1956-1957 for the Army (Jolley, 1958). Eight students were trained under the integrated concept. The experimental course length was shortened by 9 weeks and 48 flight hours, compared to the regular contact plus instrument training program. Despite this, all 8 students completed the course and qualified as Army aviators. Their contact proficiency was considered satisfactory although slightly below the regular course average. This study, it should be noted, was exploratory and was not meant to be a real test of the integrated concept. It was beset with procedural problems and difficulties. Consequently, the utility of the method remains to be established.

Summary: The available research provides no reliable evidence that giving all contact training first will have adverse effects on students

subsequently learning to fly by instruments. The issue, however, has not been examined in enough detail to permit resolution. Instrument training given early in Primary does provide skills that may become important during emergencies. It may also increase the trainee's confidence in his instruments although there seems to be no direct experimental data bearing on this point. There is little evidence to indicate that early instrument training has any adverse effects on subsequent contact flying. The apparent negative effect on night primary grades, however, does need to be clarified. It appears that, for the most part, early instrument training helps the student become at least as proficient as his control counterparts, often with less flying time, but again, no empirical evidence is available that giving instrument training first will reduce the total time required to learn to fly. The temporal course of initial advantages of early instrument training is not clear. Apparently, facilitating effects may persist at least into advanced training, since one study demonstrated that proficiency differences disappeared after approximately 200 hours of flight experience.

Research Issues:

1. Research is needed to determine if the simultaneous cue method does in fact possess significant training value. For example, systematically relating cues from two visual reference sources to each other may impose too severe a learning problem on the student pilot, with the result that overall proficiency is degraded. On the other hand, learning to use one set of cues first and then another may also degrade his proficiency when a criterion control task requires the use and integration of cues from both sources. The time-sharing hypothesis advocated by Adams and his associates (see p. 164) is, of course, relevant to this issue.

2. An additional area of research involves the investigation of learning to fly by instruments when the variable of instrument reading difficulty is controlled. Subjects given intensive familiarization in instrument reading and interpretation prior to learning to fly on instruments may well master flying fundamentals in approximately the same time as those trained only on contact "cues." If instrument reading difficulty is a major factor affecting the ease of acquiring instrument flying techniques, then the direction for change is clearly indicated.

3. Investigation of the difference in perceptual factors involved in contact versus instrument flight is also proposed. In contact flying, there are many sources of visual information and a totally different field of view than found for instrument flying conditions. Evaluation of the "newer" generation of instruments such as integrated displays, digital readouts of altitude, airspeed, etc., as factors affecting the ease or difficulty of acquiring instrument flying techniques is also implied.

4. The basic question of whether giving all contact instruction first interferes with the subsequent acquisition of instrument flying techniques remains, essentially, unresolved. The Ritchie and Michael study demonstrated that learning to fly first by contact reference makes acquisition of instrument skills more difficult, whereas the Ritchie and Hanes study demonstrated that contact-first actually reduced the time required to master instruments. Because of the obvious conflict of results of these two studies, and for additional reasons relating to instrument reading difficulty and the apparent lack of an adequate sample of representative flying tasks in earlier studes, substantial research effort in the present-day military context is needed to resolve the issue of the sequencing of instrument training.

5. As a side issue, some attempt should be made to assess the subjective factors involved in learning to fly by, or actually on, instruments. No studies were uncovered during this review that indicated whether the pilot trusted or actually had confidence in instrument flying techniques, or how confidence was best developed during training.

Light-Plane Training

Studies have been conducted to determine if training in a light plane facilitates or improves the acquisition of flying skills during subsequent military flying training. The implicit assumption underlying such training is that the trainee will acquire a degree of flying proficiency that will transfer positively to later learning. Research is reviewed here which considers the transfer of training value of such experience. Some of these studies also explore light plane flight performance for pilot selection purposes. The light plane as a selection device is discussed on page 125 of the report.

Boyle and Hagin (1953) performed an evaluation of preprimary light plane training (65 hp Aeronca "Champion") to determine its selection and training value for Air Force student pilots. Two matched groups of student pilots (n=120 each) were used in the study. Experimental group subjects were given 25 hours of light plane training during the six weeks of preflight schooling before beginning Primary Pilot Training. Control group subjects were given the standard course with no light plane training. The light plane syllabus used was specifically oriented toward teaching the basic skills important in flying the T-6 aircraft. Analysis of operational data(i.e., attrition rate, time to solo, and frequency of accidents) showed slight, nonsignificant differences in favor of the experimental group for accidents and time-to-solo (average of 4 hours less time required), and significant differences for attrition.

A subsequent study, based on more detailed analyses of the Boyle and Hagin data (Sutter, Townsend, & Ornstein, 1954) demonstrated that the light plane group was significantly superior to the control group in flying performance on objective flight checks administered at the completion of 18 hours of primary training in the T-6. On all subsequent flight checks, during Primary and Basic Training, there were no significant differences between groups. A significantly greater number of light plane students were eliminated from training prior to the 18-hour level in Primary than from the control group but the light plane group has less attrition after the 18-hour period. Overall, 87 percent of the light plane group completed T-6 training as against 62 percent of the control group. A conclusion reached was that "the principal apparent training advantage of light plane training lay in the reduction of the number of eliminations occurring during the T-6 Primary Phase."

The finding that light plane training has greater potential for selection purposes than for the enhancement of flying proficiency, has also been supported by evaluations of the Air Force ROTC Flight Indoctrination Program (FIP) (see also page 125). Preprimary light plane training (65-200 hp aircraft) apparently had no effect on subsequent primary flying performance (Cox & Mullins, 1959; Mullins & Cox, 1960).

Results similar to those of the AF FIP experience have also been reported for the Navy ROTC flight training program (Seale, 1958). NROTC men received light plane training during their senior year at the University of Illinois (n=4) and Purdue University (n=14). Upon assignment to preflight, the recipients were matched to 18 control students who had not been given the NROTC flight training. Arrangements were made to allow both groups to progress through primary at a rate dependent on the evaluation of their proficiency by the instructors or officers in charge of the training units. Consequently, neither group was restricted in progress by the number of flights stated in the standard training syllabus[12]. The light plane training students were slightly superior to the control group on several measures taken during primary flight training (e.g., primary stage A grade, number of unsatisfactory flights, number of flights required to solo). Attrition data were too limited to permit the drawing of conclusions about selection value. As none of the observed differences were

---

[12] Note that this is a deviation from the normal "hours criterion" used to train pilots. Unfortunately, no data were presented to compare these rather highly select groups with "normal" trainees under the conventional program.

significant, the conclusion was reached that "the in-training effects of the NROTC Flight Training Program do not warrant the time and money that are currently being spent on it."

Preprimary flying experience of a slightly different kind[13] was administered during 1958 to naval preflight students (Ambler & Berkshire, 1960; Berkshire & Ambler, 1963). Incoming students (n = 196) were sent through a one-week flight indoctrination phase prior to their academic training at the U.S. Naval School, Preflight. The phase consisted of four flights totaling 5.9 hours of flight time and appropriate ground instruction. Students undergoing indoctrination training exhibited slightly better flying performance (nonsignificant) in terms of primary and basic flight grades than control students. The authors note, however, that an "Hawthorne" effect (an enhancement of performance due to the knowledge that one is being observed) may have been operating to influence these grades.

Taken together, the results of the light plane studies indicate that such training is of only limited value in influencing subsequent skill acquisition and proficiency during primary and basic training. While there may be initial benefits in the form of heightened proficiency at early stages, and a small savings of training time (Boyle & Hagin, 1953; Sutter, Townsend, & Ornstein, 1954), transfer effects do not last. This is in accord with what is known about transfer of training. Most often, first-task influence is greatest on the early stages of subsequent task learning, presumably because the trainee has acquired some initial level of proficiency on the second task by virtue of the prior relevant learning. Thus, while initially he may be ahead of his control counterpart, differences in performance may tend to disappear as a function of prolonged common treatment. Whether the "disappearing differences" are due to the experimental group's slowing down (perhaps because of limits of "fidelity" of the first task) or the control group's catching up, is an open question. The Boyle and Hagin, and the Sutter, Townsend, and Ornstein studies did show significant differences in flying proficiency at very early stages in Primary but not at later stages. The Seale study (1958) and the Ambler and Berkshire (1960) study reported nonsignificant differences on various measures taken at later stages of training, in favor of students with prior plane experience. These differences might have been significant had the measures been taken earlier in training.

---

[13] The type of plane used was not identified. Since the flights occurred at NAAS Saufley, it is assumed that aircraft heavier than light planes were used.

Certainly, there do not appear to have been many deliberate efforts on the part of flight training personnel to capitalize on early proficiency gains as a result of some first-task experience, whether it be in light planes or with synthetic training devices. Lack of such efforts gives rise to the implication that training research aimed at improving pilot training methods and techniques must be directed towards discovering techniques with effects operating across the entire pilot training syllabus rather than only on portions of it. Otherwise, the proposed technique must run the risk of being rejected as not useful. It is our opinion that transfer effects of light plane training have not, in most instances, been fairly assessed, and that potential gains, where shown, have not been fully capitalized on. It is not within the scope of the present study to determine whether or not the cost of such training justifies its use.

Research Issues: It is not clear if preprimary light plane training has value for subsequent pilot skill acquisition. The data suggest that students trained in this way do exhibit slightly better initial flying performance in primary training than do control group students, but the costs of such training may not justify the apparently small gains achieved by it. Before attempting to weigh costs against gains, however, some further attempt should be made to determine precisely what these gains are. On this issue, the Air Training Command may have useful data, since the current undergraduate pilot training program begins with light plane flying. The Cessna T-41 is used as the initial training aircraft in the program and apparently satisfactory results are obtained in terms of training progression and prediction of student washout based on performance during the first 30 hours of flight instruction. It is anticipated that light plane training could be profitably used for reducing the amount of flying time required to achieve specified criteria of proficiency during primary training, particularly during the earlier stages.

Ritchie and Hanes (1964) demonstrated that first teaching the student to fly on instruments in a light plane facilitated subsequent acquisition of contact flying skills. Results obtained with the airplane were reliably superior to those obtained from a Link Trainer used in the same way. Use of the light plane as an instrument trainer prior to primary training thus may have significant training advantages for the Air Force, and evaluation should be continued.

Manipulation of Instructional Variables

Only a limited number of studies specifically concerned with variations in the methods and techniques of training pilots were discovered in the literature. Manipulation of instructional variables such as sequence of instructional units; size, composition, and complexity of these units; course content and length; scheduling of training conditions; amount of instruction, etc., can be expected to have significant effects on skill acquisition, but we were unable to locate any such researches in the aviation environment. Certainly, research manipulating such variables has been conducted for specific purposes within Air Force units, but it appears that the results have not been disseminated beyond the circle of users in the units affected (e.g., pilot training studies conducted within the Air Training Command). It is also clear that answers are not available for many questions relating to the manipulation of instruction. A recent survey (Smode & Meyer, 1966), indicated that in combat crew training schools (CCTS), the minimum/optimum requirements for training pilots are not easily nor consistently specified. Methods and evaluative data are not available for this determination. Consequently, course specification is based on judgment plus experience and expertise with previous systems as modified by the availability of time, money, and training aircraft. Program modifications are similarly achieved based on field or command information that changes are needed.

Data from educational research dealing with the effects of manipulation of instructional variables are vague and sketchy and contribute little to what is already employed in the training of pilots. Strong generalizations from such studies are not defensible; consequently, this body of information is bypassed in the present review.

To a certain extent, Air Force ground school and flying training is constrained in the sense that sequences, content, and amount of instruction are prescribed by the nature of the flying job. At present, pilots are conventionally trained to proficiency by means of an "hours" criterion. The flight syllabus (developed from previous experiences with similar aircraft and programs) establishes the content, amount, and order of training given. As conceived, the flight training programs do not permit much variation in their conduct. For example, an approach such as proficiency-based graduation, whereby students progress through training at their own rate, presents administrative problems. While the idea is technically excellent, it would, in effect, graduate students at so many different times as to place a severe load on the present administrative control of pilot output for pipeline needs.

Data on the manipulation of course and instructional variables in pilot training are scarce. Because of this, only a cursory review, based primarily on studies referring to the sequencing of training and techniques of ground school instruction, is attempted here. While studies of programmed instruction are included, the potential of this technology for pilot training can only be suggested because of the lack of evaluative data. Our emphasis is not on conclusive findings but on directions for research.

Training Variation Studies: Studies reviewed under this topic manipulated one or more aspects of the pilot training curriculum or of the training environment. An early study conducted by the Air Force Personnel and Training Research Center (Townsend & Flexman, 1954) demonstrated that changes to the then-conventional instructional techniques and methods could significantly enhance the quality of pilots graduating from the T-6 Primary Pilot Training Program. Special training methods were developed and applied to all phases of T-6 primary (ground school and flying) training. These "changes" in instructional methods produced pilots whose performance was significantly superior to that of a conventionally trained group on measures (mean error scores from check rides) taken throughout the primary training program. No deliberate attempts were made, however, to alter the content and sequence of training established by the syllabus.

Two studies were concerned with the effects of different types and combinations of earlier aircraft experience on subsequent flying performance. Johnson (1962) compared the performances of students in advanced jet training who were trained in different types and combinations of jet and propeller aircraft, and in various sequences of training in the primary and basic phases. The objectives were (a) to compare all-jet and all-prop training in both primary and basic phases of training, (b) to evaluate primary training aircraft, (c) to study the effects of eliminating primary training aircraft, (d) to evaluate basic training aircraft, and (e) to study the effects of combining jet and prop aircraft in the basic phase of training. Measures of performance taken during subsequent advanced training demonstrated that the flying performance of the all-jet group was superior (advanced flight grades) although the all-prop group was superior in advanced ground grades and completed flight training in a shorter period of time, (a) above. Differences in favor of the all-prop group, however, could be attributed to a better standardized and shorter syllabus. The evaluation of primary training, (b) above, suggested that the type of plane flown in primary had little effect upon subsequent performance. The results of the third comparison, (c) above, suggested that students can begin flight training in the basic jet trainer without

previous training in either of the two primary aircraft (the T-34, small propeller-driven aircraft, or the TT-2, small jet aircraft) and do as well in advanced training as students trained in either of the primary trainers. However, extra flying time is required during advanced training. With respect to type of basic training aircraft, (d) above, students trained in basic jets received better advanced flight grades, but the group trained in basic props completed flight training in less time (again probably due to the shorter prop course rather than to proficiency differences). Combining jet and prop training in the basic phase, (e) above, afforded no superiority over basic jet training alone. Differences in advanced training were generally favorable to students who received all of their basic training in jets.

Seale (1956) compared the performances of two differently trained groups of flight students on field carrier landing practice and carrier qualifications. Students trained in the T-34-T-28-SNJ syllabus (in a progressive order) were significantly better on some aspects of field carrier landing practice (pattern, 180° position, 90° position, final approach, cut position) than SNJ trained students, and were equally proficient on other measures. However, at a later stage of training, carrier qualifications, the SNJ students were significantly better on 90° position, final approach, cut position, landings, signals, headwork, reaction toward flight, and mental attitudes. The author conjectured that the reversal might have been due to greater experience in, and confidence with, the SNJ for the all-SNJ-trained group.

These two studies suggest that giving pilot trainees experience with a number of different airplanes during early training does not lead to better flying performance at later stages than does simply giving training in a single (relevant) aircraft. They also suggest that an easy-to-hard progression in training planes may not be necessary for sequencing pilots through training. Much more work is needed, however, to support these conclusions.

Programmed Instruction: Programmed instruction technology has value for portions of the Air Force Pilot Training Program. An air Staff policy letter[14] has emphasized the orderly transition of programmed

[14] Headquarters, USAF Policy Letter: Programmed Learning, dated 30 October 1961. Reprinted in Persselin, L. E., Programmed Instruction Technology for Ballistic Weapons Systems. BSD TDR 64-72, Ballistic Systems Division, Norton Air Force Base, California, 1964.

learning from R&D efforts into operational application in current training. For example, Air Force Manual 50-1 (Programmed Learning, 31 July 1964) encourages Air Force units to apply programmed learning to the extent considered feasible and to "continue to look for areas where programmed learning can be used to bring about efficient learning". The manual also authorizes the dissemination and exchange of programmed learning information between major air commands and operating agencies. Undergraduate pilot training programs have used programmed instruction packages for: Flight Instruments, Aerodynamics of Sink Rate, Aviation Physiology, Aviation Weather Reports, Evaluation and Measurement, Weather, Basic Navigation Equipment, Target Intelligence, Reciprocating Engines, Principles of Flight, Basic Navigation, Flight Planning, Instrument Procedures, and Radio Aids. The Strategic Air Command has applied programmed learning techniques in subject areas such as Nuclear Weapons, B-52 Aerial Gunnery, B-52 Crew Duties, B-52E and F Takeoff Planning Procedures, EWO Communications Procedures, KC-135 Pilot Emergency Procedures, KC-135 Navigator Rendezvous Procedures, and Electronic Warfare Training (Persselin, 1964). Unfortunately, no evaluative data could be located attesting to the utility and desirability of these programs over conventional instructional methods.

As a general statement, there is little available direct research data bearing on the efficiency of training pilots via programmed instruction technology. Evaluative studies of programmed instruction for other areas in human learning, however, are numerous, and reviews and summaries are plentiful (e.g., Gage, 1963; Glaser, 1964; Lumsdaine & Glaser, 1960, Lumsdaine & May, 1965; Margulies & Eigen, 1963; Schramm, 1964). Also, programming variables are discussed in Abma (1964) and principles of programmed instruction are contained in Stolurow (1965). Perhaps the most meaningful generalization that can be drawn from this body of research is that under most conditons, well-developed programmed sequences lead to proficiency on a par with that obtained from conventional instruction, and do so with appreciable reductions in training time. Martorana (1964) documented that Air Training Command experience with 60 programmed instruction packages resulted in an average time savings of 33 percent, with no loss of standards in training proficiency. One study designed specifically to investigate such reported time savings (Mayo & Longo, 1966) found that a 31 percent reduction of training time did not lead to a loss of training quality in electronics fundamentals when experimental subjects were compared with conventionally trained subjects. In the study, time reduction was included as part of the design rather than simply being an incidental finding as reported in other studies.

Some commercial airlines are currently using programmed learning techniques to teach pilot skills. Experience to date with these programs suggests appreciable time savings in transition and refresher training of expert airmen to airline standards in aspects of the pilot's job. It appears that various topics, e.g., comm/nav procedures, instrument flight procedures, safety procedures, tactics and weaponry, etc., taught conventionally in ground schools could be taught as well--and in less time--with no loss of proficiency by using programmed instruction techniques. Considerable effort may, however, be required to identify, program, and evaluate subject matter areas amenable to this form of presentation.

Research Issues:

1. An operationally important research requirement in the Air Force is to determine realistically the constitution of pilot training programs. Analytic study is needed within each command to define the validity of pilot training programs. This requires resolving such issues as: what is the optimum in course content, emphasis and training time? when is a course below minimum capability for effective training? on what basis should courses be modified? and gets at the basic issue of how pilot training programs are matched with job requirements. This capability for specifying optimum training content and time allotment for school programs is obligatory if meaning is to be derived from the manipulation of sequences of training.

2. How variation in the size and composition of instructional units and in the sequencing of course materials affects pilot skill acquisition is not precisely known. Manipulation of such variables in well-designed, complex training situations may provide valuable information on ways of improving the rate of skill acquisition and improving overall proficiency at each stage of training. Allowing students to progress through training at their own rate may be of value in other types of training situations but may not be practical for pilot training because of the administrative and instructional problems that would be generated. Consequently, "progress-at-own-rate" needs to be considered against other criteria. Research on sequence should include (a) investigation of the effects of presenting harder material first on subsequent acquisition of skills, versus an easy-to-hard sequence, and (b) determination of the effects of giving ground training on the more difficult flight control problems prior to giving instruction on the easier tasks.

3. Studies investigating the effects of number and type of training planes on subsequent proficiency should be continued. Apparently, certain types of flying experiences can be eliminated without adversely affecting subsequent flying proficiency. Research is needed to clarify this issue.

4. Programmed instruction technology has been applied to pilot training but data are unavailable for evaluating its utility. Effort should be directed toward determining which aspects of pilot training are most amenable to instruction via programmed instruction, and evaluation studies should be conducted. What this means in time and money for value received is not clear, at least to the present writers.

OPERATIONAL COMPONENTS OF THE PILOT'S JOB

The review in the preceding portion emphasized the research concerned with the efficiency of initial skill acquisition and cited important training considerations in producing proficient pilots. The research reviewed here appraises those studies investigating the performance of operationally capable pilots in critical aspects of the operational flight environment. The material which follows treats successively: low-altitude flight studies, pilot workload determination, the effects of stress on pilot performance, crew training, and visual aspects in flying.

Low-Altitude Flight

Present-day aerial warfare places a considerable requirement for low-altitude flight (from the deck to 500 feet) at high speeds (including V_{max}). The low-altitude, high-speed (LAHS) requirement is established for SAC aircraft flying Emergency War Order (EWO) missions and for TAC fighter aircraft during portions of the ground attack mission profile (interdiction, close air support) and during tactical reconnaissance missions. The demands of this flight regime will increase in severity in next-generation systems due to supersonic speeds (e.g., the F-111) and, for logistics missions, extended periods of contour flight such as will be flown in the C-5A aircraft.

LAHS flight is accomplished within an envelope defined as unsuitable (at too low an altitude) for enemy radar detection. While this reduces the aircraft's vulnerability to missiles and ground fire, it provides serious problems for the pilot, primarily increased probabilities of colliding with the ground, getting lost, and failing to find targets. A series of difficulties inhere in this type of flying. The pilot is subjected to considerable task loading due to the extra demands imposed by

controlling and maneuvering his aircraft at low altitudes (terrain following and terrain avoidance), while conducting low-level navigation and detection, recognition, and, in some instances, destruction of targets in highly stressful situations. Buffeting and gust effects, acceleration stress effects, and vibration, may be present in significant amounts. The pilot is constrained in the low-level mission context and at times is unable to perform effectively. Considerable effort is being expended to aid him in doing his job, both in terms of equipment design (displays and controls, and means for unburdening the pilot with automatic assists), and in the gathering of performance data useful for improving training. Our review of the low-altitude flight literature does not consider design research but summarizes studies on pilot performance in the air and in the simulator, as well as on training research. Since such factors as navigation (geographic orientation), target acquisition and destruction, and manual control of the aircraft have important implications for training, highlights of research findings within each of these problem areas are organized below.

Navigation: Navigating and maintaining orientation at low levels is seriously hampered by restrictions in visual reference, and the relatively high-speed flight over terrain. An analysis of the problem of geographical disorientation (McGrath & Borden, 1963) indicated that the major disturbances were caused by errors in the selection of visual checkpoints, faulty dead reckoning, and the design and use of aeronautical charts. Data based on approximately 1000 low-altitude training missions flown in light attack aircraft (Navy) indicated that geographic disorientation occurred in 27 percent of the missions flown. Army data (Thomas, 1964) indicated that aviators, during short, preplanned, nap-of-the-earth missions, exhibited disorientation in 42 out of 80 sorties. Similarly, in a Human Resources Research Office (HumRRO) field test, loss of orientation occurred in more than half of the nap-of-the-earth sorties flown, even though the sorties were only three to five miles in length.

One issue for study has concerned cartographic variables important to low-altitude flight. Recent research (McGrath, Osterhoff, & Borden, 1964; 1965) investigated the design of aeronautical charts for low-altitude VFR navigation. Two important variables identified were map clutter (e.g., place names on charts) and chart scale. These two variables in interaction significantly affected geographic orientation. Increases in chart scale alone showed no consistent effect on performance, but an increase in chart scale accompanied by a change in chart information content did affect orientation. A cinema technique was used, in which military pilots correlated map information with motion picture

information. No aircraft control or monitoring functions were required of the pilots. Unfortunately, this part-task simulation did not yield measures of performance wholly representative of performance of the activity in actual flight operations. When the method was improved, i.e., the pilot controlled the speed of the aircraft while scanning and monitoring cockpit displays during the navigation task, the authors stated that pilots "using the new method achieve better orientation performance scores than pilots using the former method" (McGrath, Osterhoff, Seltzer, & Borden, 1965, p. 40). The results did indicate however, that map characteristics influenced performance because of the speed required in orientation and checkpoint identification and in the physical act of map handling in the cockpit.

The prevalence of map reading during low-level flight was indicated in a study of Canadian Army pilots flying L-19 aircraft (Lewis, 1961; Lewis & de la Riviera, 1962). Two 100-mile sorties were flown at 100 miles per hour, 50 miles of each sortie flown at low levels (approximately 25 feet above the terrain). Film strips of the pilot's activities during the low-level portion of flight indicated that the pilot's head was in the cockpit, reading, for 27 percent of the total flying time at low level. Each map "look" ranged from 0.2 to 3.9 seconds. On several occasions, the flight became dangerous because of the pilot's preoccupation with map reading.

The high probabilities that a pilot will get lost or temporarily disoriented during low-level flight emphasizes the need for training in low-level navigation, search, and use of navigational aids.

Target Acquisition: Visual acquisition of targets and release of weaponry under LAHS flight conditions is a formidable task because of the complex visual requirements generated by the high rate of closure between aircraft and object. In a short time span the pilot is required to perform visual search, assimilate the presented information, detect and identify the target, maneuver the aircraft, and initiate strategy for taking action against the target. Evidence indicates less than desired inflight performance in this difficult portion of flight. In an Air Force study conducted at Eglin Air Force Base, Florida (cited in Miller, 1964), pilots flying current TAC fighter aircraft at low altitudes and at speeds ranging from 350 to 700 miles per hour were required, in one pass, to detect, recognize, and release ordnance on target jeeps and tanks. In only 10 percent of the flights was the goal successfully accomplished.

In another Eglin AFB test involving acquisition and identification of targets (Wade, 1964), a sample of 30 pilots flying at 500 feet altitude and at an airspeed of 475 knots were asked if during the test runs they could have made some type of munitions delivery on the targets. In only 32 percent of the passes did the pilots feel that weapon delivery could have been made.

Similarly, data obtained by Joska (1965) at Eglin AFB suggested that the speed of flight significantly affected both target acquisition distance and target acquisition probability. Eleven pilots flying F-100 aircraft were to acquire 2 1/2-ton trucks at 500 feet of altitude at speeds varying from 250 to 550 knots. Each proceeded into the designated target areas and upon target acquisition, notified a radar site by calling "Tallyho." The point of acquisition was marked on an x-y plot and the pilot called out the target symbol as he passed over it. Each pilot made four passes over different targets at the preplanned speeds of 100-knot increments from the 250-knot speed. Although the data are minimal, they suggest that the pilot's capability to acquire ground targets decreases dramatically as his speed increases from 250 knots to 550 knots. The implication is that at still higher speeds target acquisition may not be achieved. No attempt was made to camouflage or otherwise confound the visual targets.

These findings are important in that they provide operational flight data (albeit sketchy) on visual acquisition of targets. More data obtained in the operational environment are needed to permit comparisons with laboratory findings. While laboratory study provides the most control and a high reliability in performance prediction, it is well known that predictions from the laboratory to the real world are inexact. The above findings, for example, would not be predicted from the analytical (laboratory) data from visual research. The analytical data generally provide performance values overly optimistic of the values obtained from the field data. Perhaps the analytical information best describes the upper boundary of performance rather than typical human behavior in flight.

Snyder (1964) has identified parameters which exert great influence on visual performance in detecting and recognizing individual ground features, and has indicated the extent of current knowledge, mostly from laboratory research, of how each parameter affects dynamic visual performance. Two classes of parameters are indentified which most significantly affect air-to-ground visual performance. These are, the physical and geometric, and man's visual capabilities. In the first class, the parameters include: the masking of the target by terrain features; the

size and shape of targets; illuminance, luminance, and contrast; clutter in the visual field; and time available for search (groundspeed, etc). The effects of human visual limitations must consider object luminance and contrast, angular velocity and blur, and static versus dynamic acuity. These parameters are discussed in some detail by Snyder. The point is clearly made, however, that for a systematic data generation program in which the parameters of target size and shape, contrast, illumination, terrain features, angular velocity, etc., and their interactions, are investigated, operational field data are needed with which laboratory data can be compared.

Control of the Aircraft: In addition to the visual-perceptual requirements in LAHS flight, control of the aircraft's position and altitude in relation to the ground is a demanding aspect of the pilot's job. The important variables in this flight situation include terrain features, airspeed, navigation requirements; and vibration, buffeting, and gust effects. Particular concern has been voiced about overloading of the pilot in LAHS flight, and much research activity has centered on the understanding of buffet, gust, and vibration effects on pilot performance. A number of controlled laboratory experiments have been performed to determine the effects of vibration on humans. These have investigated physiological effects, subjective tolerance levels, mechanical body responses, and aspects of performance as related to frequency and acceleration (or amplitude) of vibration. Vibrations above 20 cps on extended missions generally cause only discomfort and fatigue but do not result in performance decrement. Most laboratory studies investigate vibrations below 20 cps, for these occur most frequently in flight and effects detrimental to performance have been demonstrated. Impedance studies have provided information about (1) body resonances at 4 to 6 cps and 10 to 14 cps, (2) head resonance at 20 to 40 cps, and (3) thorax-abdomen resonance at 3 to 4 cps. However, differential effects result due to body position and restraint systems used (see Coermann, Magid, & Lange, 1962; Goldman & VonGierke, 1960).

Tracking studies which attempted to identify frequencies, amplitudes, or accelerations at which performance is impaired, have not yielded any consistent usable data. It is clear that performance decrement is not related solely to increases in frequency, amplitude, or acceleration. The research indicates that between 1 and 30 cps, impairment is related to amplitude, acceleration, and frequency, but in a way difficult to specify. Greatest performance decrement appears between 4 and 8 cps, depending on amplitude, and may occur above 10 cps when visual acuity is impaired. Fraser, Hoover, and Ashe (1961) found that vibration in the z- and x-axes hinders performance more than vibration

in the y-axis, although the effects of vibration on multiple axes have not been determined.

Experimental laboratory studies have provided only minimal information of value for pilot training. The data are sparse and deal with simplified task characteristics not nearly approximating the complex of variables and their interactions found in LAHS flight. Unfortunately, inflight programs at present are so hampered by instrumentation difficulties, unpredictable environmental conditions, problems in experimental control, and the hazards of this type of flight that it appears the inflight environment is better suited for operational data collection to support simulation research than it is for controlled experiments. Simulation is the most useful technique available for systematic research on the variables and their interaction in LAHS flight.

Few studies have been specifically interested in assessing pilot performance during simulated, low-altitude, high-speed flight. Most notable are the North American Aviation research studies (Rawson, 1963; Soliday & Schohan, 1964; and Soliday, 1965) which utilized a dynamic flight simulator (G seat). The device consisted of a vertically moving cockpit with a total travel of 12 feet and a capability for accelerating up to ± 6 G, a functional control system and cockpit display, and an analog computer for obtaining solutions to the equations of motion. The pilots flew simulated LAHS terrain-following missions (continuous maintenance of 500 feet clearance above ground) with varying task loads defined by terrain ruggedness, navigation requirements (heading changes), airspeeds ranging from .4 mach to .9 mach, gust intensities ranging from 2 feet/second to 10 feet/second, and a number of introduced emergencies. These studies indicated that the pilot in these simulated flights (light fighter jet aircraft, advanced-type surveillance aircraft, and a jet possessing F-111 characteristics were mechanized on the computer) can fly LAHS terrain-following missions effectively throughout a broad envelope of task loadings without crashing or exceeding 1000 feet altitude (arbitrarily defined as "missile kill" height). Steepness of terrain slopes had the greatest influence on terrain-following performance. Difficulty in maintaining the 500 feet terrain clearance increased in direct proportion to increasing slope steepness. Airspeed was the second most important variable affecting performance. The indication is that pilots can track well in a medium heavy gust environment at .7 mach but tracking deteriorates rapidly when airspeed reaches .9 mach. Missions could be flown in turbulence producing acceleration loadings as high as .4 RMS G on the pilot. Variations in navigational requirements (making heading changes) and in emergency tasks (e.g., hydraulic system failure, pitch augmentation failure, engine fire, etc.) appeared to have no affect on

altitude holding ability (Soliday, 1965). Vertical acceleration and tracking did not induce sufficient fatigue to affect performance for periods up to three hours in severe acceleration environments (Soliday, 1965). Similarly, no fatigue effects or physiological disturbances under the G and task loadings were noted by Soliday and Schohan (1964) during 90-minute flights. Heart and respiratory rate were in the normal range and biochemical tests revealed no organ damage. Rawson (1963) indicated that three hours of continuous buffeting can be tolerated but the risk of incapacitating fatigue is high.

The studies indicated marked learning effects occurring under LAHS flight conditions. Schohan[15] suggested that pilots in the simulator learned "tricks of the trade" in muscular adaptation to turbulence and G's, in dial reading, in adjusting the restraint system, etc. Reduction of fear and a feeling of competency occurred when pilots experienced such flights in the simulator.

Perhaps the most significant difficulty in studies of this sort is the issue of precisely how performance from the laboratory/ simulator can be predicted to the real world of LAHS flight. Heavy reliance has been placed on simulators and laboratory research to obtain human performance data having implication for training and training research. While the studies are, in large measure, sophisticated and well conducted, they suffer from artificiality of conditions. For example, Sadoff and Wempe (1965) cite research on the effects of vibratory acceleration stress on performance in a terrain-following task. The task was performed inflight and also in a moving cockpit simulator for varying levels of turbulence. Performance in the simulator (RMS altitude error) held up to a level of about 0.3 G RMS, while flight results showed considerably more scatter and up to a threefold increase in error. The difference between simulator and flight performance were in part ascribed to greater pilot workload in the air (no lateral disturbances or navigation tasks were included in the simulator) and to the absence of hazard in the simulator. The data were based on 20-minute runs both in the air and in the simulator. The effects of stress and anxiety in the flight situation, and the complexity of the job structure constrain somewhat the prediction of performance based on laboratory data. While this is a problem besetting laboratory research in general, it is of particular meaning in this situation due to the severity of the operational flight environment.

---

[15] Personal Communication

Correspondingly, there is at present, a lack of data obtained from operational LAHS flights. The absence of good flight data has hampered the assessment of the validity of simulator data.

As indicated earlier, inflight data, though meager, suggest that a compound of factors conspires to minimize the pilot's ability to detect targets and navigate while flying at these perilous altitude-speed combinations. It is obvious that task overloading is an issue of some concern, and the desire to unburden and assist the pilot has been the impetus for much of the research activity (e.g., design of automatic systems, display and control systems for terrain following and terrain avoidance, etc.). Thus, the "help" the pilot requires has come, and will come, from improvements in equipment design. It must also come from an emphasis on training in critical aspects of low-altitude, high-speed flight.

The important areas in LAHS flight are well known to the workers in this field and were enumerated at the conference on problems related to low-altitude flight, mentioned earlier (Miller, 1964). Those parameters of specific importance to this study are listed below.

> The role of the pilot in the system must be defined precisely. This is important for mission analysis and for designing automatic systems to augment the pilot's capabilities.
>
> The pilot must be unburdened during the low-level portion of the mission. This can be achieved by improvements in displays or eliminating certain tasks which he currently performs. The result of this should allow the pilot to minimize in-cockpit viewing with more time for extra-cockpit activities.
>
> Adequate maps for low-altitude flight and methods of map display must be developed.
>
> Improved training procedures for low-altitude, high-speed flight are needed. In addition, criteria should be established for methods of visual navigation.
>
> There is a need for a closer integration of simulation research and operational studies in order to define and use more realistic parameters.

There is a lack of data obtained under operational flight conditions. These data are needed to determine how meaningful laboratory results are in the low-altitude, high-speed flight context.

Turbulence and gusts and the aircraft bending mode appear to be more important problems in low-altitude, high-speed flight than does vibration.

An important problem area needing further study is the visual acquisition and identification of checkpoints, particularly, the pilot's ability to estimate range, angles, rates, etc. These are critical in decisions on when to override an automatic terrain following system.

Training Research: Little in the way of training research has been directed specifically at parameters of LAHS flight. The studies accomplished have centered on visual-perceptual efficiency in aerial observation and target recognition.

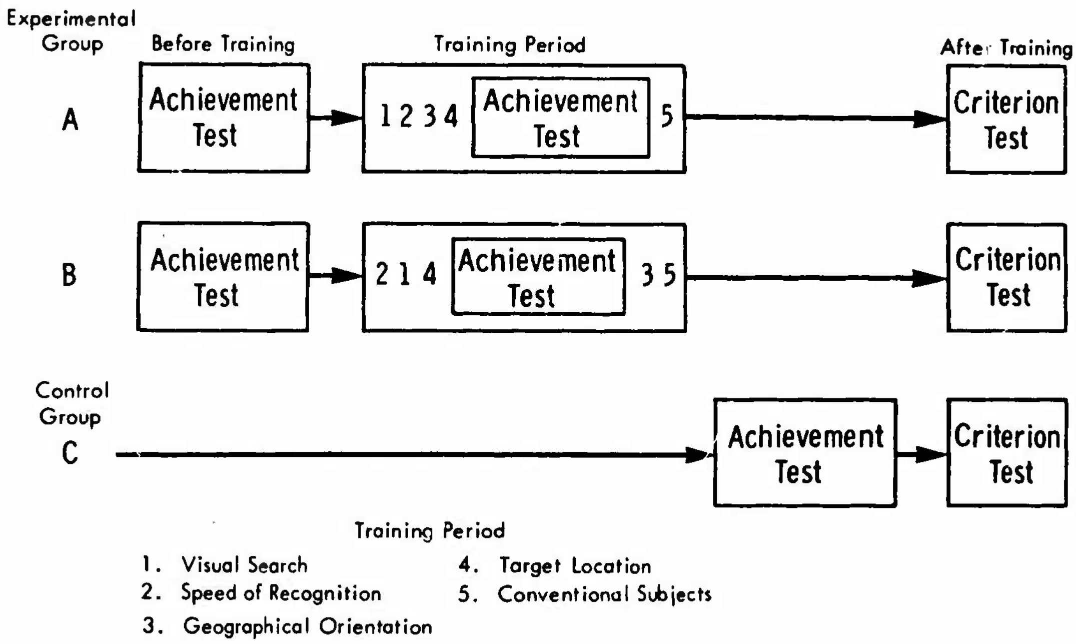
The purpose of one study in the Eglin AFB series (Wade, 1964) was to determine the pilot's perceptual efficiency during LAHS flight runs as a function of the method of ground training given in target recognition. Two groups, each of 15 Air Force pilots matched on aircraft flown and experience, received training in target recognition prior to actual flight test. The experimental group received training on a tachistoscopic teaching machine, while the control group received training by studying picture booklets. Performance of the experimental group after five training sessions (5 1/2 hours total) was significantly higher than the control group in identifying 18 targets in a 72-photograph presentation. Each target was projected on a screen for one second. Four different views per target were used: e.g., close and distant ground, and close and distant aerial oblique views. One curious feature, however, was that all 72 photographs were used in training by the experimental group whereas the control group trained with only 26 from the total of 72 photographs, the remainder being different. Following the ground training each pilot flew four sorties along a defined flight path at 500 feet altitude and 475 knots airspeed in either an RB-66, RF-101, F-100, or F-4C aircraft. Four different targets, a jeep, a 2 1/2-ton truck, a tank retriever, and a "weasel," were used. No significant differences in performance between the groups were found for the following measures: (1) number of successful target identifications, (2) measured target acquisition distances, (3) measured target identification distances, (4) elapsed

time between target acquisition and identification, (5) the pilot's estimations of target acquisition and identification distances. The conclusion was that the tachistoscopic teaching machine training did not differentiate between the two groups. This is justifiable for the conditions of the study. It appears that the level of training was not sufficient to demonstrate proficiency in this complex skill; hence, the value of the tachistoscopic technique was not fully assessed. A program of training is required which fully utilizes the technique in training to a precise criterion. In other words, considerable improvement in these skills may be needed before significant differentiation in target recognition occurs.

The Human Resources Research Office has conducted a research program for developing methods for training personnel in low-altitude aerial observation. The initial effort was devoted to identifying the skills involved in a future low-altitude task environment. From this, a tactically realistic field test was devised and four skill areas were identified: target detection by visual search; quick target recognition; geographic orientation; and target location. Based on this information, five field experiments were designed to develop methods for training observers in these identified skill areas, and achievement tests were developed for evaluating the effectiveness of the training methods (Thomas & Caro, 1962). In essence, the experimental results suggested the following:

> Performance of students given recognition training was no more accurate than that of students not so trained in reporting military objects not included in the training. Thus, intensive and extensive recognition training is required to satisfy system requirements.
>
> Maintenance of geographical orientation requires adequate observer aids.
>
> Side movement search is a most effective search method for systematic visual coverage.
>
> There is a substantial relationship between the student's ability to maintain geographical orientation and his accuracy in target location. The question to be answered is whether or not the magnitude of the location error can be significantly reduced by appropriate training.

Training methods developed from this study were incorporated into an experimental course of instruction (Hesson & Thomas, 1962; Thomas, 1962). Three groups of 18 observers each, two experimental and one control, were identified. The two experimental groups received 32 hours of training: 18 hours of experimental instruction, 3 hours of flight (at speeds of 80 to 100 mph), and 11 hours of conventional training. Officers in the control group received the minimal observer training as specified by Army directive, which included 78 hours of classroom instruction and 20 hours of practical flight. This group also had a median of 19 hours experience. The design of the experimental training is shown below. The achievement test (2 forms) was a practice mission in target location and required integration of the skills involved in training. The criterion test simulated the kinds of conditions predicted for future battle area observation and was made up of 27 target groups located over an 80-mile flight path. Both tests were scored for accuracy of object identification and for map location of each target complex.



The results of this study indicated that a systematic training program emphasizing visual perceptual skill development provided for more efficient training in the basic skills of low-level observation than did the current conventional method. Comparable learning can be accomplished in a much reduced time period.

Based on the above research, a programmed text was developed to teach skills in the four basic content areas of low-altitude aerial observation (Dawkins, 1964). The teaching efficiency of this course was comparable to the experimental course cited above, as evaluated by a paper and pencil criterion test which was substituted for the inflight measure. HumRRO concluded that the programmed course of instruction is a practical method for teaching low-altitude observational skill, and offers the field commander a convenient means of training students without heavy involvement of instructors or aircraft.

The training implications from the slowly accumulating data on pilot performance in LAHS flight are obvious--pilot training programs should place a greater emphasis on LAHS flight. This includes the development of precise performance criteria, measures, and measuring techniques associated with positioning the aircraft, navigation accuracy, and target-detection accuracy.

Research Issues:

1. It appears that system performance during LAHS flight can be enhanced considerably through improvements in equipment design (e.g., improved inertial navigation systems, improved terrain-avoidance radar, improved displays, etc.). An essential research requirement is to examine the LAHS flight regime to determine how best to spend research money on improving displays and other equipment to facilitate pilot performance. System improvement through improved hardware may reduce considerably the training requirements in the LAHS mode.

2. A serious problem in low-altitude, high-speed flying is that of maintaining geographical orientation with respect to the target. Inabilities to acquire a target, to navigate adequately or to deliver ordnance on a target compromise mission effectiveness. The results of laboratory research and field data strongly suggest the need for continuing effort on the important visual parameters in low-level flight, defined earlier. Simulation studies should investigate visual acquisition and identification of targets or checkpoints as a function of speed and altitude combinations (e.g., the pilot's ability to estimate rate of change, angles, range, perception of objects under various terrain conditions and camouflage, etc.). In addition, operational data from field investigations are required for validity assessment. The need for the study of human capabilities and tolerances in the laboratory is quite clear, but the complexity and stresses of LAHS flight make it doubtful that laboratory data are wholly meaningful in the operational context. The symposium on low-altitude flight reported by Miller (1964) stipulated that operational data must be

collected in order to evaluate pilot performance and permit comparisons with laboratory data. For example, operational data are needed to validate the mathematical derivations of target characteristics. The laboratory studies are too clean in terms of real-world visual aspects of targets, hence provide inexact data.

3. Training research could be devoted profitably to the development of course materials for target recognition (see Thomas, 1962). Emphasis on visual-perceptual training employing materials suggestive of low-level contour flying with targets ultimately placed in a natural ground milieu is desirable. Programmed degradation of the visual scene should be considered as well as time compression (including tachistoscopic presentations). The development of methods for visual search to be used during flight (e.g., time sharing, orientation strategies, etc.) could also be part of this effort.

4. The disheartening results of the published field tests suggest the need for a low-altitude, target-detection, ground-orientation training device. A relatively simple indoctrination and practice device seems indicated for use as an adjunct to visual perceptual training in checkpoint identification, navigation strategies, target location, etc. A device is envisaged which is essentially static, requiring a continuous scene presentation, a cockpit mockup, and a discrete response capability.

5. A systematic program is indicated to determine those simulation factors that provide transfer of training to operational low-altitude, high-speed flight. Continuation of the study of motion effects on transfer is certainly desirable. The available literature suggests that motion cues in the simulator enhance transfer of training. Earlier studies (Muckler, et al., 1959; Townsend, 1956) indicate that motion simulation is desirable particularly in early maneuver training. Simulating motion provides the trainee with additional cues compared to the static condition and this facilitates performance. This was affirmed by Besco (1961) in a study of tracking in the pitch dimension in simulated terrain contour flight, and by Feddersen (1961) in a simulator study of hovering training in a helicopter. The research issue is: what effects do turbulence motion (gusts, buffeting) and acceleration stresses have on training for LAHS flight? There is evidence that the greater initial transfer of training with motion, compared to the static condition, dissipates quickly. Feddersen (1961) found that his training group performed better initially in the air than did the statically trained group but the difference disappeared by the end of the six-trial flying sessions. The greater initial transfer may be explained simply in terms of familarity with a larger number of task aspects. Two other studies using the Grumman simulator

for training and criterion tasks (Buckhout, et al., 1963; Ruocco, et al., 1965) reached similar conclusions. Although these two studies lack experimental precision, the facilitating effects of vertical turbulence motion cues were demonstrated (see also page 153).

6. The role of stress in low-level, high-speed flight is not adequately understood. It is a serious issue, however, when one considers the hazards of flight; fatigue; the effects of buffeting gusts, vibration, acceleration; and heat. A better understanding of the stress components in flight is indicated.

7. One aspect of LAHS flight that has received little attention concerns the motivation of pilots for hazardous flight. Data are sparse on the courses of action pilots would follow and their expectation in difficult flight environments (e.g., LAHS in nuclear combat). A study by Jones and Lindsey (1965) investigated the attitudes of TAC fighter pilots (number of subjects ranged from 55 to 65) concerning their ability and the ability of their fellow pilots to perform LAHS flight, and their estimates on the lowest altitudes they could comfortably maintain under certain conditions. Also investigated were the relationships between personal factors (anxiety level, flying experience), and estimates of lowest altitudes maintainable. The findings indicated that stress and fatigue are major factors in low-altitude flying. The pilot knows when he is operating under great stress but does not know when he is reaching the limits of ability. Other indications were that these pilots evidenced confidence in their ability for LAHS flight and would fly at lower altitudes for short time periods (10 minutes) at slower speeds (below 300 knots) when the conditions of terrain, turbulence, and visibility were most favorable. Significant differences existed among pilots on the minimum altitudes they would fly. Those grouped as high in anxiety level uniformly gave higher altitude estimates than those grouped as low in anxiety level. This suggests a relationship between personality factors and the manner in which a flight may be accomplished and indicates the use of these data for pilot selection. This line of inquiry suggests the need for further research effort in the area of motivational factors in hazardous flight.

Pilot Workload

In the design of equipment for operators in man-machine systems, frequent use is made of workload analysis to determine the physical and mental effort required in task performance. For example, research on flight control system design or display-control compatibility would be interested in the percentage of normal working capacity the pilot expends in task activities under the given experimental conditions. This informa-

tion would serve usefully in the making of decisions on design optimization. Workload analysis is also useful in the development of procedures for pilot training. It is obvious that the pilot is more heavily loaded in job requirements in certain portions of the mission profile, for example, during LAHS flight, weapon delivery phases, and recovery from emergencies. In certain emergency situations, the combination of vehicle dynamics and environmental conditions in interaction with the speed of failure onset and the complexity of procedures to be remembered for the emergency, may overload the pilot. Thus, the important issues for training center on the capacity of the pilot to handle the job requirements, the extent to which additional tasks or job requirements can be handled (e. g., what happens when the pilot is given too much to accomplish during a stage of training), and what role learning and experience play in job accomplishment. Estimates of pilot workload, then, could serve usefully in the development of operational and training procedures, providing data supportive to that obtained from stress research, and the study of the content and the sequencing of training.

Pilot task loading is most conventionally viewed as an equipment problem, the issue being one of unburdening the pilot in the loop. In the LAHS flight context (Miller, 1964) the question is resolved into determining which tasks can be eliminated from the pilot's purview. By using the pilot strictly in those tasks at which he excels and providing automatic assists in other areas, performance decrements resulting from operational overloading should diminish (reduction of workload is correlated with the degree of equipment automaticity). In order that the dichotomy between manual and automatic be precise, an earlier stated need reappears, i. e., the pilot's role in any proposed flight mission environment must be known and defined in detail based on the proposition that he is an active element in a closed loop servosystem.

Not much empirical data is available on workload estimates for operator tasks, much less for the pilot's job. How to determine the amount and meaning of task loading and overloading has been difficult for research and the evidence is cursory. We were unable to obtain any quantitative data on workload for mission tasks that were valid for training; hence, the review centers on available methods for workload analysis. Information handling capacity has been the locus of concern in the investigations on pilot workload. The logic is this: since there is a continuous interchange of information between the pilot and his aircraft, inputs and outputs can be described and measured to such a quantitative degree that information processing rates can be defined. In essence, the comparison of this with known (or predicted) maximum information processing rates provides a measure of pilot workload. Thus, to obtain esti-

mates of pilot workload, the total time required by the pilot to perform the flying tasks is obtained (i.e., the sum of input, processing, and output times). This is then compared with the time available to complete the system tasks. The former information is based on actual obtained measures; the latter is based on systems analysis data (requirements and performance criteria). Obviously, analysis of workload is more easily accomplished for discrete tasks or discontinuous control tasks than it is for continuous control tasks. Just as obvious is the realization that describing pilot workload is a tall order.

Siegel and his associates developed a technique for applying information theory concepts to analyzing pilot activity (in an APOLLO program report submitted to Minneapolis-Honeywell Company, cited by Cole, et al., 1963). Pilot activities were partitioned into the smallest sequentially ordered tasks. Each subtask completion time was computed and all subtask values summed to give a completion time for the activity. This was compared with time available for the activity to obtain an estimate of pilot workload. Display complexity was measured in bits ($H = \log_2 n$) where n = number of equiprobable alternative readings of the display). Information processing time was computed from the formula:

$$\gamma = a + \hat{b}\hat{H}$$

where:

γ =reaction time (seconds)

a =lower limit of human response (known as 0.2 seconds)

$\hat{b}$ =reciprocal of information handling rate

$\hat{H}$ =amount of information in the display (in bits)

Pilot output times, i.e., the times required to put information into the system, were obtained from methods-time measurement.

Determining workload in continuous control tasks is a more subtle undertaking. In order to assess the operator load, the more artful techniques require the subject to perform secondary (auxiliary) tasks while simultaneously performing the primary task. If secondary task performance is good, it is assumed the primary task is easy (i.e., task does not impose a substantial workload on the operator); if performance is poor, then the primary task is defined as demanding since the assumption is that there is a limit to the rate at which an individual can handle information. When the limit is exceeded, errors occur. The secondary task may also be used as a stressor for the primary task. This was the use made by Garvey and Taylor (1959). Two tracking systems were employed, both operated equally well in terms of tracking error scores but differed

in the degree of effort required. Under the stress of loading tasks such as warning light monitoring, mental arithmetic, etc., performance decrements widened between the two systems.

The secondary task is also used to determine how much additional work the operator can accomplish while still performing the primary task satisfactorily. In the measurement of pilot workloads during performance in two alternative X-15 aircraft control modes. Ekstrom (cited by Knowles, 1963) employed a secondary self-paced pushbutton task which was performed in conjunction with the primary control task. A matrix of 16 touchlights was placed peripherally to the subject's primary task. As a light came on, the pilot extinguished it by depressing the light. Scores from the secondary task were converted into an operator loading index to demonstrate differences in primary task difficulty.

The primary task workload (W) was expressed as:

$$W = 100 - (W_1 + W_t)$$

where:

W = primary task workload

W_1 = loading task workload

W_t = eye transition time workload

W_1 was computed from

$$W_1 = \frac{N}{N\,max} \times 100$$

N = number of lights handled per second

$N\,max$ = calibrated maximum number of lights per second

W_t was computed from

$$W_t = \frac{\text{transitions}}{\text{total time}} \times 100 \times 0.14$$

(0.14 = average transition time)

The index yielded the percent of the pilot's effort (attention) that was devoted to the control task.

Knowles (1963) had identified a number of characteristics that the secondary task should possess. The task should not interfere with primary task performance, be simple (require little learning), be self-pacing, be compatible with the primary task, and be easily scored. Tasks capable of meeting these criteria include the previously mentioned one of attending to lights placed peripheral to the operator, mental arithmetic, monitoring, and self-adaptive tracking.

The display intermittency technique is another method used to measure pilot workload. In this technique the displays used by the pilot are obscured for some length of time. The percent of time the displays are obscured is systematically varied, the assumption being that the proportion of the time displays must be present is a more sensitive indicator of performance than measurements of error under normal conditions. Using this technique, Lindquist and Gross (1958) studied vehicle control in a single axis. The task was to center a dot on an oscilloscope center line with a conventional stick grip, score being represented by the time integral of dot displacement over the problem period. The dot was made to disappear for increasingly longer intervals until control accuracy was lost. The viewing time (i.e., dot visible) expressed as a portion of the total time was defined as the workload of the manual control task.

The secondary task technique provides an overall indication of operator loads. It also gives an indication of how load varies during a task and permits an assessment of learning during the acquisition of the primary task, e.g., when the primary task performance does not change and the secondary task performance becomes progressively better, this is evidence of increasing mastery of the primary task.

It is clear that pilot capability is governed by information handling capacity as well as the momentary potential of the pilot to react (physically and mentally). A major variable in the capability is amount of training and experience. With overlearning of the flying tasks and repeated experience with handling emergencies, the pilot becomes more efficient in selecting, coding, and translating information, which results in better use of, or greater, channel capacity.

As mentioned earlier, present-day training considers the problem of workload tacitly. Emphasis is placed on increasing the pilot's skill level and providing him experience with overload conditions. For example, he receives much training in aircraft emergency procedures. Thus, improvement in skills and experience with emergencies increases the workload capacity of the pilot. The emphasis is towards overlearning which results in such benefits as: reduction in mediating responses (with decreases in reaction time, effort, and fatigue), localization of response (dropping out of extraneous response components), increased anticipation of event happenings, reduction in habit interference, increased accuracy and reliability of response, increased learning of contexts, and self-confidence.

Research Issues:

1. A number of problems concerning the workload of the pilot can be profitably examined. A prime issue ties in with recommendations made previously, namely that the role of the pilot in current mission environments must be defined quantitatively and in detail. One way to attain closure with this difficult requirement is to study pilot activity in information exchange terms. There is a continuous exchange of information between the pilot and the aircraft and many of the pilot inputs and outputs can be described in quantitative terms. Information processing rates can be computed and information handling capacity defined. Comparision of actual information processing rates for the varieties of sequential and coordinative activities in flying can be compared with known and predicted rates for a better understanding of pilot workload. Certainly, such results would have important implications for equipment design, especially in situations where there is a requirement to unburden the pilot, where feasible, in order to achieve his optimum integration into the dynamic control system. However, such data are also of use in research aimed at gathering information that provides a base for predicting the endurance capability of the pilot in the mission context, and ultimately should be of use in specifying training requirements.

2. A continuing effort is needed in the development of methods for analyzing pilot workload. The present rudimentary techniques, using motion-time analysis for assessing workload in discrete tasks in systems, and the methods employed in laboratory tracking studies, such as secondary tasks, stress techniques, display intermittency techniques, physiological measures, and frequency measures, could serve as the rallying point. The methods should be addressed to assessing workload in the operational mission context. Better measures for quantifying workload are sorely needed. These should be developed to indicate, ultimately, performance capacity per level of skill development.

Performance and Stress

The disruptive effects on performance resulting from perceived stresses is of some consequence to aviation, yet, formal sequences for coping with stressful events are not part of the pilot training curriculum. This may be due, in part, to the difficulty of defining stress operationally and, in part, to an inability to institute procedures for exposing trainees to stress events. Few research data on the effects of stress in the pilot's job have been generated to provide support for training. It is possible that operational personnel subscribe to the view that training for

stress is best accomplished by providing experience in flying and practice in specified emergency procedures.

The topic of stress, however, is well represented in the literature both as to its concepts, its conditions, and its effects on human behavior. A review of the literature indicates that stress is a multidimensional event not easily conceptualized but readily discussed. It is regarded as having both external and internal sources, as an intervening variable (with emotional, motivational properties), or as a stimulus condition. The conditions producing stress (both short- and long-term causes) have been identified variously as task- and environment-induced (overloading, underloading, unexpected stimuli; emergency, noise, heat, shock, etc.), and failure-induced (harrassment, personal loss, too-high standards, failures, etc.). Stress is also correlated with physiological factors (secretions, changes in body functions), social factors (conflict, frustration, anxiety), and with personality. That this body of literature is well represented is confirmed by the reviews or summaries prepared by Chiles (1957), Deese (1962), Klier and Linskey (1960), Lazarus, Deese, and Osler (1952), and others.

Much of this literature deals with the conditions and effects of actual stresses in laboratory-contrived situations and is only of passing interest here. Our concern centers on the operational implications of stressful events on pilot performance. The important feature is individual response to stress rather than the causes and conditions of stress, i.e., how the pilot reacts to stressful situations and what happens when stress is introduced in a training situation. Observation and anecdotal information indicate that substantial stress on the pilot increases muscular output and tension, with a loss in coordination. Fatigue and sterotyped behavior may occur as well as temporal and spatial narrowing of the perceptual field. Of extreme importance to flying is the tendency for fixation and loss of short-term memory (e.g., remembering instructions in holding, landing patterns) as well as cognitive loss (e.g., thinking, planning, programming of operational sequences).

Very little in the literature deals directly with stress training in aviationlike environments, primarily because of the difficulties involved in installing realistic stress situations. The laboratory studies have typically attempted to simulate stressful events with such stimuli as electric shock, noise, fatigue, unpredictability of consequences, and information overload. These approaches all fall short to some degree in representing the stresses found in flying (e.g., fear of physical harm). Unfortunately, in the laboratory experiment, subjects quickly develop a "set" that no harm can befall them in the study; hence, their behavior is

oriented to the experimenter's expectations (i.e., "Act as if I'm frightened"). Thus, the situation is unreal, as are the resulting physiological and behavioral measures. Berkun, Bialek, Kern, and Yagi (1962) have labeled this denial of threat in experimental studies as "cognitive defense" and describe it as the principle obstacle to the study of human response to stress. Similarly, studies using achievement failure or threats to ego integrity, while stressful, are different from real-world stresses associated with fear of injury or death. The problem of structuring an experimental task in which the stress component is realistic and meaningful to pilots is perhaps the most difficult in the whole domain of behavioral research. On the one hand, the threatening event must be physically safe and must not suggest psychological harm to the subject afterwards. On the other hand, the event must be simulated in a credible way by overcoming the cognitive defense mentioned above.

Experience with training cadets to fly has shown that fear and anxiety are importantly related to success in the program. While anxiety symptoms are demonstrated by most trainees, significant correlations between failure rate and heightened anxiety have been observed over many classes in military flying schools. An example of this long-term concern is the study by the Naval School of Aviation Medicine (now the Naval Aerospace Medical Institute) (Voas, Bair, & Ambler, 1955) which investigated the relationship between behavior in a minature stress situation and manifest fear in flight training. The stress situation was a simulated high-altitude flight in the decompression chamber. A total of 1540 cadets were taken to a simulated altitude of 20,000 feet and instructed to remove their oxygen masks. High-anxiety students were defined as those who replaced their masks during the 10-minute stay at altitude and/or complained of ear block (pain due to pressure on eardrum) during descent. Anxiety reactions in the chamber were significantly related to later anxiety toward flying resulting in dropout from the program. This work, part of an effort to improve selection of naval aviators, yielded results which suggested the possibility of developing measures to serve as screening devices or as criteria for tests of stress tolerance. Investigation of the miniature stress rationale has been continued (see Longo & Doll, 1962) but the studies apparently did not get beyond the feasibility stage.

Most of the stress studies have employed an actual stress event and assessed its effect on performance. A limited number of studies have not used an existing stressor but have examined the effects of anticipatory stress on performance. Anticipating an unpleasant social or ego damaging event, for example, can disrupt behavior noticeably (see Deese & Lazarus, 1952; Lazarus & Baker, 1957).

Two programs of research within this latter frame explored the anticipatory threat of physical harm. This type of threat (injury or death) is of considerable relevance to pilot training, for it may temporarily, but seriously, disrupt flight performance. The first of these programs was conducted by the Human Resources Research Office largely in a nonaviation environment. The second program, undertaken by the Naval Aerospace Medical Institute, explored the development of methods for identifying stress-susceptible individuals, a goal being the selection of potential pilots not overly susceptible to the threat of physical harm. Because of their relevance to pilot training, both programs are discussed in some detail.

The first program, a series of imaginative studies by HumRRO (reported by Berkun, 1964; Berkun, Bialek, Kern, and Yagi, 1962) investigated the causes of behavioral degradation under psychological stress, and attempted to develop procedures for reducing the severity of th. problem. Beginning with the basic assumption that any man will break down if exposed long enough to the stresses of threat, a purpose of the research was to develop training procedures for retarding deterioration in performance resulting from stress. Emphasis was placed on achieving a stressful situation experimentally by recreating those elements of naturally occurring disasters that have a fear effect. To do this requires that the threat perceived by the subject be induced by cognitive stimuli, i.e., the subject has to think about the whole situation and figure out that he is in trouble. Thus, the subject is provided information of such realism that an assessment of various events leads him to believe that he is actually being threatened. Determining the effectiveness of performance under stresses is based on a stratagem of cognitive stress stimuli (threat acceptance by the subject) and the ability to measure objectively the performance relevant to the stressful situation. This type of task situation removes much of the artificiality of most laboratory research (i.e., the subject's knowledge that he will not be harmed, disbelief in the fidelity of task simulation, and a task-taking set).

Three criterion elements were chosen to determine operationally that an effect is involved like that elicited by a naturally occurring threat.

1. Subjective Self-Report--experimental subjects should demonstrate a significant negative affect, defined in these studies, by a list of quantitatively scaled words.

2. Performance Measures--the distribution of scores made by the experimental subjects must differ signigicantly in location or shape from the control group. Various task-relevant measures involving

speed, accuracy, and completeness were obtained, as well as a consolidated measure (composite performance score) which combined a relative speed score with a pass-fail score for the subtasks.

3. Physiological Responses--the experimental subjects should evidence a disruption of normal physiological processes comparable in kind and intensity to that found in a naturally occurring threat or in combat. Two measures were obtained. The first was a urinary sample to determine 17-hydrooxycorticosteroid accumulation. The second was a sample of blood from the finger for a count of circulating eosinophil cells. (Eosinopenia, which is a decrease in the number of circulating eosinophils, has been demonstrated as a sign of "alarm response.")

Five stressor task situations were developed and employed in the program of research and compared with control situations on each of the above classes of measures. The tasks are described below.

Situation 1: The subject is threatened with injury or death and he cannot actively resolve the predicament.

Ditching--The subjects were passengers in an aircraft which was in "trouble" and was preparing to ditch or crashland. All overheard a pilot-to-tower conversation concerning the emergency and could see crash equipment on the airstrip. These were the supports to the deception.

Situation 2: The subject is threatened with injury or loss of life but is able to do something about it.

CBR Warfare--During a maneuver, the subject, stationed alone at an isolated outpost is required to radio reports to the command post on the presence of aircraft overhead. He later hears over his radio set that a nuclear accident will result in a dangerous fallout of radioactive material in his area. Immediate rescue is possible if the subject can report his position over his radio which has suddenly gone dead. The maneuver is canceled and all activity is now concerned with evacuation of personnel from the area. Perceptual confirmation of the hazard is provided by a radiation dosimeter available at the position. In order to be rescued, the soldier must repair the failed radio.

Forest Fire--The setting is the same as above except that the "accident" is a forest fire surrounding the lone subject's outpost. Perceptual support is provided by artificial smoke generated nearby. A failed radio interferes with rescue and it must be repaired by the soldier.

Artillery Shell Barrage--A series of explosions simulates an artillery barrage and the subject hears, via radio, that the barrage has gone astray and shells are hitting outside of the designated target area. He sees also that the shells are falling in a pattern which will hit his position. The explosions constitute the perceptual support. The subject's transmitter inexplicably fails, although he continues to receive messages. Rescue depends on repair of the transmitter.

Situation 3: The subject is not threatened with injury but is made to feel responsible for an injury to a buddy.

Demolition Explosion--The subject, as part of a work detail setting up a training problem is instructed in wiring-in explosives placed in a canyon below. Working alone, the subject is instructed to match colored wires with colored wires already on screw posts and upon completion, throw a switch, which will then enable others in the canyon to use the circuit. Immediately on throwing the switch, a 5-pound charge of TNT goes off in the canyon. The subject is then informed of a man being injured in the accidental blast which may have resulted from incorrect wiring. The subject is instructed to telephone Fort Ord, but the telephone does not work and his calls over the intercom are ignored, which makes it appear that he cannot be heard. The subject, however, receives a variety of messages over the intercom for the next 35 minutes, asking about his progress in calling Fort Ord, that the military police want to question him, etc., and also hears urgent messages concerned with keeping the injured man alive.

In each of the 5 experimental situations, the number of subjects exposed ranged from 13 to 27. Three of the five situations met the criteria of stressfullness as defined earlier. The most extreme results (stressor task) were obtained in the situation where the subject believed he was responsible for an explosion which injured another soldier. The simulated aircraft emergency aloft and the artillery shell barrage, both threatening the subject's life, also satisfied the three criteria that the stimulus complex produces an effect similar to that evoked by naturally occurring threats. The CBR task satisfied only the subjective self-report and the physiological steroids-level criteria, whereas, the forest fire task group differed significantly only on the steroids-level criterion.

These studies yielded a significant result, namely, that a stimulus complex can be installed which simulates the stress effects elicited by naturally occurring threats. Such an "apparently real" approach permits the meaningful study and assessment of various stress levels on operationally performed events. Care must be exercised, however, neither

to violate ethical considerations in dealing with humans nor to invoke psychological damage that is a residual of the experimentation. The concepts underlying the research represent a step beyond the less than plausible laboratory task situations and the effects achieved. The possibilities for training research and the eventual development of techniques and procedures for retarding the onset of stress effects in the flight environment are excellent.

In this vein, a program of research is underway at the Human Resources Research Office, Division 6 (Aviation), Fort Rucker, Alabama, to assess the decrements in performance resulting from various stresses associated with aviation combat missions.[16] At present, information on human reactions to stresses in flight missions is being gathered and examined, including anecdotal information from Vietnam operations. However, little data exist which describe behavior changes of a pilot in combat.

An interesting sidelight to the concern for inflight observation is the Ames Flight Research Center development of instruments and methods for quantifying the normal physiological responses in man to the stresses of flight. One output of this effort is a portable system that monitors and records an individual's heart rate. During late 1965, a team of three NASA scientists gathered heart rate data from Navy carrier pilots during actual combat operations in Vietnam.[17] Heart rate and vertical acceleration records, obtained from the time a pilot entered the cockpit to touchdown back aboard the carrier, showed that heart rates were higher during the periods just prior to takeoff and landing than they were during bombing runs. These results are consistent with other data obtained on NASA and USAF Aerospace Research Pilot School pilots, which show that risk (for experienced airmen) has a negligible effect on heart rate but responsibility causes large changes. A specific example is the study of Roman (1965) which was part of an inflight monitoring program conducted at the NASA Flight Research Center. Two pilots in a high-performance aircraft (F-104B) were instrumented during a number of demanding flights. It

---

[16] Communication from Dr. Wallace Prophet and Dr. Wiley Boyle.

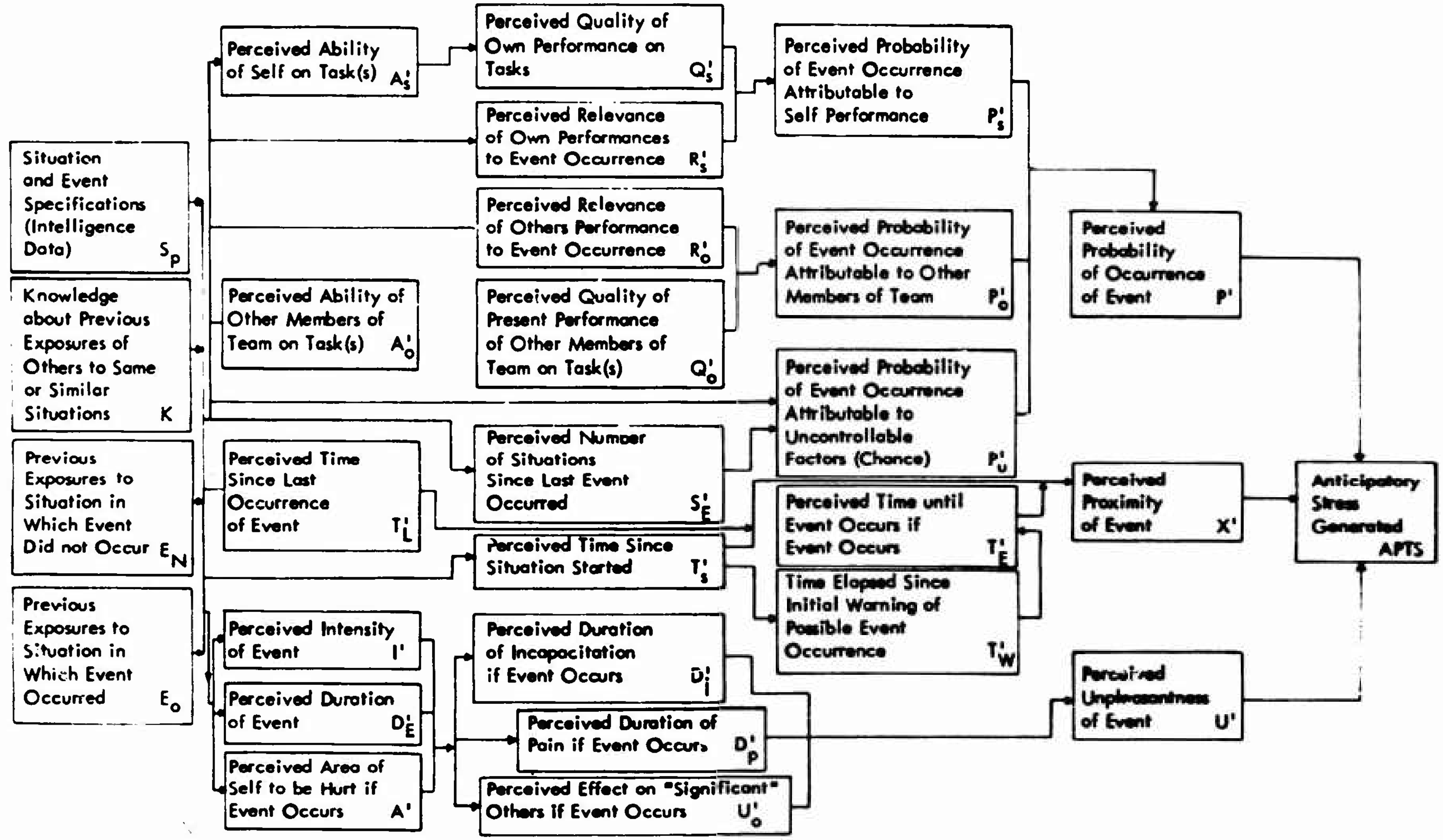
[17] Congressional briefing given during 1966 by W. L. Jones, Director, Biotechnology and Human Research, Office of Advanced Research and Technology, National Aeronautics and Space Administration.

was found consistently that the pilot actually flying the aircraft showed a much higher heart rate than the pilot not in control. This was also true when the roles of the two pilots were reversed. It was concluded that responsibility for the mission, or responsibility coupled with risk appears to be a much more potent factor than pure risk in producing high heart rates. The suggestion that the body does not react in a direct manner to the presence of stress (or risk) alone is quite significant for the interpretation of biomedical information.

Concurrent with the inflight observation, a conceptualization scheme is being developed by Division 6 (Aviation) for handling the available information and determining critical information gaps. Based on these findings, experimental research of two types is proposed. The first consists of determining parameters of the effects of stress on operational performance. The other centers on studies of training to reduce or eliminate performance decrements due to stress. The eventual research pursued will, in part, be determined by the outcomes of the initial data-gathering phase.

It appears that this effort could fruitfully explore the development of confidence-building techniques as a means of stress retardation. One research lead concerns stressing the communications during the tactical portion of flight (e.g., jamming, heavy traffic in landing zone). Another possibility is the investigation of conditioning techniques to instill habits useful under stress. For example, there is some evidence that helicopter pilots, when fired upon, will respond by pulling the aircraft up (vertical response), thus, hover. The response is not useful in this situation since the opposite is desired, namely, getting out of the area fast.

The second program of interest relevant to pilot training is the Naval Aerospace Medical Institute research on the development of methods for the early identification of stress-susceptible individuals. A study by Wherry and Curran (1965) explored the possible psychological variables contributing to the generation of anticipatory physical threat stress (APTS). A model was developed in which the major determiners of APTS were hypothesized as (1) the relative unpleasantness of the event (U'), (2) its proximity (X'), and (3) its probability of occurrence (P'). The varieties of evaluations the individual must make for each of these determiners in the model are shown in Figure 5. An objective of the research at this time was to validate portions of the model so that situations analogous to the model relationships could be installed to define individual stress susceptibility. Obtained differences could then be systematically related to later behavior in the aircraft.



(From Wherry & Curran, 1965)

Figure 5. Determiners of Psychological Stress.

Electric shock was used as the threatening event for the experimental study since it is credibly threatening but not physically or psychologically damaging to the subject. Sixty-four cadet pilot trainees were subjected to a four-choice color discrimination task. This simple task was characterized as "information processing in a simulated aircraft emergency." As each stimulus was presented, a corresponding response key was depressed by the subject at a self-paced rate. The length of each trial ("mission") was shown by a row of lights which were consecutively lighted for 10 seconds each, down to time zero. Shock, either mild or painful, was administered only at time zero according to two probability-of-occurrence levels (.2 and .8). The procedure was as follows. First, the task was explained (X') and the electric shock demonstrated. Second, the subjects were familiarized with the probability generator and that the occurrence of shock at time zero was determined randomly (the generator was a metallic drum whose surface was divided into a metallic portion and nonconductive tape portion on which moved a metal stylus). Third, the shock level, U' (mild or painful), and the P' values (.2 and .8) were set by the experimenter, the probability generator started, and the session begun.

The initial results suggest that P', U', and X' are major determiners of anticipatory physical threat stress, combining in a multiplicative fashion. Disruption is greater when the unpleasant event gets closer in time, when the perceived probability that the unpleasant event will occur is high, and when the perceived degree of unpleasantness is increased. Also, behavior in subsequent stressful situations is influenced by previous experience with the unpleasant event. Although APTS is a hypothetical construct and never measured directly in the study, there are indications that it has a curvilinear relationship to performance with low amounts of threat enhancing performance. The conclusion of importance is that the amount of APTS in a situation can be controlled and its effects studied in the laboratory.

Research Issues:

1. Research can profitably examine techniques for reducing emotional anxiety reactions that might impair performance. Emphasis should be placed on the conditions that evoke stress responses and on the development of procedures for reducing anxiety-provoking potentialities. Assuming that performance degradation will eventually occur if the individual is subjected to a stressful situation long enough, a desirable product of the research would be the development of training procedures for retarding the onset and the severity of the effects of stress. The training sequences should emphasize providing experience with the

determiners of psychological stress, and hence, building up situational confidence in flight.

To achieve this, a laboratory program of research is needed on the determinants of effectiveness of performance under the stresses encountered in flying. The direction initiated in the studies of anticipatory physical threat (Berkun, et al., 1962; Wherry & Curran, 1965) seems highly appropriate in concept for this pilot training research. Of special interest is the development of situations (relevant to flight) which induce the trainee cognitively to accept the simulated threat as real. The need is for a cognitive stress situation involving fear (or guilt) and on which objective measures of performance can be obtained on relevant tasks.

2. There is also a need to obtain data on pilot performance during operational flight, i.e., handling emergencies in the air (and in combat situations) so that the value of laboratory results can be assessed. We cannot rely completely on laboratory simulation to determine the effects of stresses on performance. Objective measures of pilot performance in flight under real-life conditions (stress and anxiety) are needed. Recent refinements in miniaturi~~zation~~ of sensors, transducers, and recording equipment make such inflight data collection entirely feasible.

Crew Training

The interdependence of human behavior is a prevailing feature in military flying, and effective interaction between airmen is highly desired. Teamwork or coordination is a frequent requirement in this structured task-oriented job situation. Not only is coordination a requirement in multicrew aircraft, e.g., Strategic Air Command, Military Airlift Command, but also in single- and two-place aircraft such as flown by the Tactical Air Command and the Air Defense Command where pilots perform in multiship elements, often in interaction with ground control. Thus, in defined instances, the effects arising from the interdependence of behavior influences aircrew performance and accordingly are of consequence for training. The requirement for training is not only to enhance the capability of the crew to perform according to formalized standard operating procedures and to handle contingencies as they arise, but also to enhance performance through coordination of crew activities.

Aircrew training objectives are derived from the nature of the missions flown and the job requirements in flight operations. Thus, conceptions of crew behavior and interaction are constrained by the feature that individuals perform in a structured job situation wherein each mem-

ber has a defined function, interdependent with those of other crewmembers. This position orientation makes explicit the tasks assigned each crewmember and the sequences of performance required to carry out the specifically defined missions tasks.

Our review of the team training literature is primarily directed at research in the above format, i.e., studies in behavioral integration in position-oriented, somewhat structured job situations characteristic of aircrew operations. Data are specifically sought on the ability of the pilot to coordinate the effort of the crew to achieve mission objectives and the capability of the crew to receive, process, and act on information generated from various sources in the time-shared, forced-pace mission environment.

An enormous number of studies are available concerned with human behavior in the group context, and our initial task was to select pertinent studies from this research. Since very few studies are directly relevant to pilot training, care was exercised in delimiting the range of studies reviewed by selecting only those researches that had some demonstrable relationship to flying tasks, even though remote or even conjectural. Immediately eliminated were the studies of human behavior in small groups concerned with effective socialization of group members, e.g., researches in group dynamics, interpersonal relations, role playing, and similar social-psychological interactions. These groups are usually informal in structure and rely on the independent contributions of individuals, most often in the framework of problem-solving, decision-making activities. They have little if any resemblance to aircrew structure.

A number of laboratory studies have investigated the functional roles of humans in relatively simple task situations and in simulations of components of complex systems. For the most part, these studies are also of minimal significance to pilot training, and coverage of these findings is well beyond the scope of this report. We have, however, selected a number of representative studies from this classification for review where the conclusions have general implications for crew training. Studies of team training in the context of large military systems are also sampled. While these situations are quite unlike the aircrew environment, the concepts and findings are nevertheless of interest to crew training and suggest important hypotheses for training research. The most relevant studies are those that investigate the behavior of aircrew in synthetic environments and during flight, and are described here in some detail. Unfortunately, the studies are few and the data available are sketchy.

Laboratory Study of Team Performance: Considerable experimental research has attempted the systematic examination of the variables which influence team behavior. Utilizing relatively simple laboratory tasks, several classes of variables have been manipulated: task load and patterns of work assignment, information feedback, and task differences on team organization. While these classes of studies yield findings of theoretical significance, the results are, for the most part, not directly relevant to pilot training and hence are not included here. Summaries of parts of this literature have already been prepared in some detail. (See, for example, Alexander & Cooperband, 1965; Moore, 1961). Certain programs of research, however, offer results that suggest concepts and hypotheses useful for training research and hence, selected studies are outlined below to indicate the nature and specificity of the data.

Roby and Lanzetta (1956; 1957) and Lanzetta and Roby (1957) investigated different conditions of communication structure on small teams by controlling the pattern and sources of information given the subject. Three-man teams were installed with two controls and two instrument displays at each position. Each subject made control settings based on direct instrument readings and also on readings received from other team members. Communication structure differed in the degree to which subjects had direct access to needed information. Autonomy of function defined the extent to which information was available at a position. Under low autonomy, each member was dependent on the others for information; under high autonomy each subject had all the needed information. As predicted, differences in team performance resulted as a function of task communication structure. More errors were made when a member depended on another for relevant information and errors increased when both members had to provide information. In the low autonomy situation, the stress of time pressure actually disrupted team performance. The team could not process information as rapidly as an individual could, and it was difficult for the team to set up channels and organize for processing information.

The composition of information feedback on problem solving and information processing was the issue in a program of study by Hall (1957). The task of each subject in a two-man team was to turn a knob, the combined movements of both members determining the placement of a pointer on a micrometer scale. Each subject was provided individual feedback on the adequacy of his own response, and feedback (confounded) was given to the team about how well they were coordinating to produce correct adjustments of the pointer. Teams trained under confounded feedback improved more quickly in accuracy than when trained under individual feedback. The reason was that subjects learned to compensate

for each other's errors. The confounded feedback condition provided higher role differentiation to subjects, since they had access to knowledge on the appropriate weighting for their control with respect to the teammate. When another condition was added which provided only feedback of the teammate's response (Rosenberg & Hall, 1958), this resulted in the poorest team and individual accuracy. The suggestion is that team performance is the result of the performances of the individuals that make it up. Feedback based on the performance of the whole team may distort the information each member receives concerning the adequacy of his own behavior.

Glaser and his associates (Egerman, Klaus, & Glaser, 1962; Glaser & Klaus, 1962; Glaser, Klaus, & Egerman, 1962) conducted a program concerned with manipulating team proficiency by applying operant conditioning techniques to the acquisition and extinction of team responses. Considering the team as the module of investigation, proficiency was made to vary by controlling the reinforcement supplied to the team. In these studies, the task of the individual team members was to depress a switch on a panel for either 2 or 4 seconds, depending on a pattern of lights displayed. Two members (monitors) made the response independently, and a third member (operator) made a response based on his judgment of the correctness of the two monitors. If both monitors make the correct response, the operator produced a team reinforcement, which registered as a point on a counter visible to the team. Thus, reinforcement occurred only when all individual responses were correct. This confounded feedback defines the condition where the team is reinforced by a single event, the occurrence of which depends on the integrated responding of the members on any one trial, i.e., the group feedback is contingent upon a composite of individual performance. The findings from these studies suggest that the teams have response features which are directly affected by the feedback from team output. Team acquisition is a direct function of the conditions and schedule of team reinforcement during team training as determined by the probability of a correct team response.

In a search for optimum methods of team training, a group led by Horrocks became interested in defining the kinds of skills needed for effective team performance. In one study, (Horrocks, Krug, & Heermann, 1960) six-man teams performed the activity of decoding jumbled sentences. Two team members (decoders) received words sent by another member (router). A fourth member (evaluator) completed the task of word decoding and two team members (integrators) arranged the words into sentences. Team training involved interactive performance among all members, whereas in individual training the members practiced their

tasks separately. The results indicated that team practice did not enhance the learning of individual skills. In the same study, four-man teams, composed of three estimators and an integrator, determined, from a set of 10" by 10" stimulus cards, ground range and azimuth to a spot beneath a simulated aerial target, based on elevation, azimuth, and range estimations. Different methods of providing knowledge of results were utilized. The outcomes indicated that crude feedback (direction of error but not magnitude) resulted in better performance than did more precise feedback. The authors concluded that too much information feedback cannot be used effectively in the situation where the input is variable and where little improvement occurs over time. Directional information was all that was needed.

In another study (Horrocks, Heermann, & Krug, 1961), the effects of experience as a working team on performance was investigated. Several task situations were employed (decoding cryptograms, a paper and pencil matrix comparison test, and constructing five-letter words). The results of these studies indicated that team performance was most dependent upon individual skills. Coordination, or group enhancement effect, emerged only as the result of high levels of individual proficiency. For example, a member of an intact functioning team may be replaced by another equally competent person without detriment to team efficiency. It was also found to be unnecessary for team members to receive their initial training as a unit. Team practice did not enhance individual performance. Initial training in team performance proved to be more efficient when conducted on an individual basis. Evidence also suggested that if coordination is emphasized during early phases of training, it will interfere with the acquisition of individual skills. The authors make it clear that these findings are limited to highly structured teams performing relatively simple laboratory tasks in the acquisition phase of learning.

Findings suggesting the superiority of individual training over team training were also obtained in studies of team performance in a simulated radar control of an air intercept task (Briggs & Naylor, 1965; Naylor & Briggs, 1965). The transfer task utilized the Ohio State Air Traffic Control Simulator (Hixon, et al., 1954). Each member of a three-man team was provided a CRT display on which appeared eight target and eight interceptor radar returns which moved in real time. The requirement was to achieve intercepts. The training task involved a checkerboard with eight target and eight interceptor checkers. The results of Naylor and Briggs indicated that team performance was influenced by the way the task was organized. Teams in which individuals worked independently performed better than teams that interacted heavily in problem solution, for example, in exchanging targets and interceptors, and in

verbal communications. In fact, training in the team context encouraged communications habits that inhibited team performance in the transfer task.

The relevance of the task situations exemplified in the above studies to the pilot's tasks are tenuous, at best. The interesting feature, though, relates to the conceptualization of team interaction. An outcome of value from these studies is that team performance is heightened through the development of individual skills; coordination emerges as a result of high levels of individual proficiency. Johnston (1966) argues similarly that an emphasis on individual training is more useful in team activities than is training in the team context. He lists several factors in support of this. First, individual skills are more essential at transfer if team members perform largely independent functions, particularly when available training time is limited. Second, individual skills may be more difficult to learn, whereas required team skills may already be in the trainee's response repertoire. Third, team skills may be acquired in the context of individual training. Fourth, team training may encourage habits which retard team performance in the transfer task. Another outcome from the studies just cited is that team organization and operations, and procedures for team training, are determined by the nature of the work.

Quite opposite to the foregoing is an approach developed in the context of large computerized command/control systems training which considers team operations in an environment characterized by emergent situations. In an emergent situation, all action relevant to environmental conditions has not been specified, and the state of the system is not related to standard procedures or relied-upon predictions (Boguslaw & Porter, 1963). Organization and procedures are <u>developed by</u> the team rather than <u>imposed on</u> the team. Although position and task assignments are defined, the team member is permitted a degree of latitude in performing in terms of the contingencies that arise in the system, and the team as whole adapts to the emerging characteristics of the environment. Proficient team performance depends upon cognitive organization of the environment. Thus, proficient teams may vary considerably in operating procedures, relying on procedures that have come to be appropriate to the particular team, and coordination is dependent on the development of plans which integrate the peculiar operating procedures of all members. This approach has been most artfully developed by the System Development Corporation (SDC) beginning with the system training program in the air defense context (see Alexander & Cooperband, 1965; Boguslaw & Porter, 1963; Goodwin, 1957; and Kennedy, 1962). For a number of years SDC has been involved in developing management

procedures for Air Force personnel engaged in operating modern electronic defense systems as coordinated teams. A research that represents this effort in improving team performance in the emergent task situation is the air defense experiments of the Rand Corporation, (Chapman, et al., 1959; Kennedy, 1962). Teams ranging from 28 to 40 men were placed in a simulated Air Defense Direction Center and given the assignment of defending a given area against attack by air. These complex teams, performing information processing and command and control functions, comprised a surveillance section, movements and identification section, and a weapons control section. Early warning stations, aircraft, an adjacent direction center (handover of aircraft), and higher headquarters completed the man-machine system. In essence, the job was to detect, acquire, and identify moving aircraft (tracks). If a track conformed to a known flight plan, the aircraft was designated as "friendly"; if it did not match, it was designated as "unknown" and interceptor aircraft were dispatched to determine if the intruder was hostile.

The richness of the findings in this long-term research program on organizational behavior cannot be adequately presented here. Simply stated, however, it was amply demonstrated that team coordination develops and improves in an unstructured way without guidance or directive. Teams may be equivalent in overall performance yet differ considerably in procedure, since performance depends on how the environment is organized cognitively by the members. Several equally good solutions to problems usually exist, and interteam differences in procedures become evident in coping with the environment. The development of coordinative skills is stressed, but adequate individual skills are required. The task of the trainers is to install conditions which will enhance the development of team integration. Alexander and Cooperband (1965) suggest three concepts underlying team development: development of system awareness, development of an integrated model of the environment, and the exploitation of the potential self-organizing capability of the team. A set of principles of team development has been outlined in general form by Kennedy (1962) and is listed below.

> The job of the manager is to develop people who are capable of achieving organizational goals or mission of the system.
>
> The manager should treat and deal with the crew as an organism or whole, not in terms of individuals, during the growth period.

The manager should insure that the synthetic organism gets the information that it needs in order to develop adequate channels for a flow of information inside the synthetic organism dealing with task accomplishment, and should arrange reinforcements so that all components of the organism are aware of successes and failures in achieving the organizational goals. "Debriefings," involving the whole crew, after each period of operation help to perform this function.

Reinforcement should be factual and impersonal. It should deal with the problem of accomplishment of the goal of the organization, not with individual performance.

Motivation to achieve the goal of the organization should not be exhortative, but should be given frequently through communicating the value of achieving the organizational goals.

The manager should develop an attitude in the organization that encourages and rewards "invention" at all levels.

Individual performance should be assessed and rewarded or punished in terms of contribution to the organization's goals.

Training for effective team performance involves additional considerations in addition to individual skill requirements. A number of these, described by Boguslaw and Porter (1963), are summarized below.

Orientation to team goals: emphasis on understanding the consequences of error to the team output.

Training in interdependencies: interdependent relations between and among team members should be incorporated into the training curriculum.

Training for error analysis: skill in analyzing one's own errors is of extreme importance, and the process of learning to deal with errors in the team situation is of high priority. Problem solving discussions rather than didactic presentations is one way of encouraging the problem solving attitude.

Training for sensing overload: operators must learn when overloading is imminent and when to seek help; also when a team-mate is nearing the overload condition by watching for the occurrence of errors for which they can take compensatory action.

Training in adjustment mechanisms: under overload conditions adjustments of various types can be made in that the team may operate so as to permit queuing, engage in omissions, commit errors, filter some inputs, approximate the exact content of messages sent, increase the number of channels of work flow, or reduce the number of categories in any classification activity. Application of such concepts in the training curriculum is useful in preparing individuals to withstand greater periods of task overloading, and to meet the demands of emergent situations more easily.

Training for emergent situations: on-the-job training is needed to instruct individuals to sense problems created by changes in the system environment.

This system training is quite different from aircrew training, both as to complexity and objectives. Usually, system training is reserved for later sequences of training when system hardware is well defined and when individual skills have been learned. By design, its goal is to train teams to perform functions difficult to practice in any other integrated manner. The features of emergent situations, partial reliance on formal operating procedures, impromptu response invention, etc., however, have meaning for aircrew training and are discussed next.

Aircrew Training: A problem underlying crew training and research is the inability to define what is meant by a "good crew" or coordinated crew behavior. This has made it difficult to install training situations and devise measures descriptive of crew behavior. It is well understood that crew interaction occurs in aircraft where positions are highly structured in terms of responsibilities and SOP, with tasks and task activities specified. Relatively formal operating procedures and communications exist and a wide range of behaviors is required, with the results of performance comparable to objective references. Yet, views of crew interaction and concepts of coordination vary depending on the referent situation and the assumptions made concerning crew behaviors. Using the language employed in computerized command/control systems, two classes of events occur: established, and emergent. The former are repetitive and predictable, with specified and detailed rules for handling them. The latter are unpredictable and may have more than one equally good solution. Thus, coordination, at times, results naturally from a sequence of properly planned and executed individual acts. Individual skills, it is argued, are the learned components, while coordination, or group enhancement effect, emerges only as a result of high levels of individual proficiency. Within this framework coordination refers to synchronized team action involving mechanical coordination by means of formalized standard operating crew procedures. Activity is timebound in that the cues initiating behavior come either from the completion of activities by other members or from time signals. In other words, crew effectiveness in these routinized task situations is seen as the sum total of the individual performances. Coordination, however, also results when members interactively perform in situations where there are no predetermined standards of performance. This hypothesis emphasizes improvisation and impromptu response invention. In training, one would present task situations that had not been practiced to such an extent that performance had become routine, and the emphasis would be placed on the adaptive innovations developed by the team members. In short, effective performance is here regarded as something more than the summation of individual skills (Krumm, 1960). The established and the emergent situation responses operate in a complementary manner in aircrew activities.

The difficulty in conceptualizing team behavior, and the differences concerning what to observe, have generated problems in assessing crew output. Development of relevant and reliable crew performance has not been encouraging. Attempts at obtaining end-product measures of crew performance have not yielded good, stable results. In the B-29 crew research (Forgays & Irwin, 1952) objective indicators of crew performance on 600 crews in training showed poor reliability for such indices as

circular error scores, camera scores on simulated visual bombing missions, average errors in making good control times, etc. As a result of this, the early crew research program was forced to utilize ratings made by superiors since these possessed a certain authority status and demonstrated better reliability (Sells, 1958).

The effects of social factors on bomber crew performance have been of interest to the Air Force, and studies have attempted to establish relationships between attitude and performance. A study of 89 B-29 crews flying combat missions in the Far East (DeGaugh & Knoell, 1954) found significant relationships between combat performance and dimensions of crew attitudes. The highest correlation was between "pride in the workgroup" factor and superiors' ratings. The items subsumed under this factor referred to the sense of trust in, and liking for, crewmembers, and crew professionalism. A moderately high correlation was found between the "job satisfaction" factor and superiors' ratings. Another study of B-29 crews (Knoell, 1956) also demonstrated a correlation between crew attitudes and rated combat performance. The interpersonal factors of pride in workgroup and crew acceptance of the same task-oriented values were related significantly to the criterion.

Some study has been devoted to assessing performance values deriving from the rational assembly of aircrews. The supposition is that some combinations of personalities, backgrounds, skills, and other characteristics will enhance crew performance more than others. Haythorn (1957) has summarized a number of researches which support the contention that crew effectiveness is in part determined by variations in combinations of individuals which make up the group. The importance of this as a research issue has been underscored by Sells (1958), who advocates the development of proficiency measures and personality test profiles as well as mathematical models as techniques for the study of problems in crew assembly. The difficulty with this sort of evaluation however, is the supposition that satisfaction with the crew situation, pride in group, high morale, and other concepts denoting "happiness" is indicative of good crew performance. The impressive evidence from industrial psychology is quite clear in showing that performance and satisfaction systems are not consistently related and often are negatively correlated. Fiedler (1958) is one of many who have demonstrated this. In his studies, low-achieving B-29 crews were concerned more with social satisfaction and pleasant individual interactions whereas high-achieving crews were concerned with skill and competence in task performance. While crew feelings of well-being and enjoyment of group participation may come from skilled crew performance, it is not a condition of successful performance.

Crew coordination may be regarded as latent ability, that is to say, it is a potential of crews to respond effectively when some unusual or demanding circumstance arises. Since coordination tasks in aircrew are often not in a face-to-face situation, one attempt at describing behavior has been through examining crew communications. Initial studies in the Air Force began in the early 1950s to determine if crew conference techniques could be used to develop methods for enhancing crew coordination.[18] The Crew Operating Procedures (COP) test was developed for evaluating the level of coordination in crews. This work, with RB-47 crews, was not successful however, in defining criteria for evaluating crew coordination training (see Hood, 1960). The use of the COP test as a possible criterion measure of crew coordination was continued in a study of B-52 crews with the expectation that it might serve as a criterion measure in evaluating performance in the simulator. Technical problems with the test, however, precluded its use (see Krumm, 1960).

Research has suggested the possibility that crew effectiveness may be gauged by the amount of time the individuals spend interacting. Thus, measures such as volume, content, and patterning of communications appear promising as an index of coordination potential. Crew coordination studies were conducted in the B-52 integrated simulator facility located at Castle AFB, California. The integrated trainer is a B-52 flight simulator (MB-41) electronically linked to an APQ-T2A Ultrasonic Bomb-Nav Trainer. The device is actually a subteam trainer for two pilots and two navigators functioning as a crew. A study by Krumm and Farina (1962) attempted to assess the training value of this device. The purpose was to investigate the effects of training crews together in the integrated configuration as compared to individual crew training. The value of the linkage device was determined by comparing the experimental group (n = 38 crews) with a control group (n = 37 crews) in accomplishing the final integrated mission requirements. Communication measures were developed as one means of defining crew coordination. Measures were taken of volume (production of message units) and pattern (kinds of message units such as voluntary inputs, acknowledgements, etc.). Based on the results of the communications analysis, simulator flight checks, and a crew evaluation questionnaire, the integrated simulator sequences promoted crew coordination. The communications scores, however, were not unequivocally acceptable as a criterion measure.

---

[18] Credit for this approach goes to I. K. Cohen who observed that the "Eagle" crews undergoing B-29 training for the Korean War at Forbes AFB, Kansas, developed their own procedures for interaction where SOPs were not well defined.

An evaluation was made of the integrated simulator as a means for improving the proficiency of experienced B-52 crews (a study cited by Siskel, Lane, Powe, & Flexman, 1965, portions of which are classified). Extensive recordings were made of crew interphone communications during Emergency War Order (EWO) type sorties flown in the simulator. Number of transmissions per minute and number and kinds of messages transmitted per minute during 30-minute bomb runs were computed. Bomb-run performance was also assessed by 16 objective performance measures. Experienced crew performance improved over a sequence of four sorties with a concomitant decline in crew communication. A comparison on the final sortie of crews trained in the integrated sequences with control crews, experienced but not given integrated training, revealed that the experimental crews performed significantly better in terms of the measures taken. An inverse relationship existed between communications scores and the other objective performance scores. These results suggested the possibility of developing an objective measure of crew proficiency based on the inverse relationship between communications and performance. Consequently, Siskel, Lane, Powe, and Flexman investigated the communication processes in B-52 and KC-135 crews of differing levels of experience during segments of peacetime mission profiles. The hypothesis was that more experienced crews would have lower communications rates (transmissions and messages) than less experienced crews. Eleven B-52 crews (six combat and five student crews) flew a ten-hour aerial mission which included aerial refueling, navigation legs, and several high- and low-altitude bomb runs on a radar bomb/scoring site. The mission segments chosen for analysis were the takeoff and the bomb run. Ten KC-135 crews (five combat and five student crews) flew a six-hour sortie which included aerial refueling and navigation legs. The mission segments scored were the takeoff and aerial refueling. Crew transmission and message rates were obtained from tape recordings of interphone communications. Based on the results of this sample, the hypothesis was not confirmed. Some problems were experienced in experimental control during the flight missions.

Research Issues:

1. The emphasis on integrated crew training has shifted somewhat because of reductions in crew composition for aircraft programmed to become part of the inventory. The F-111, for example, will have a crew of two; the massive XB-70, assuming an operational configuration, will require a crew of four (2 pilots, defensive systems operator, and offensive systems operator). Crews for advanced flight vehicles are similarly envisaged to be small. Thus, intracrew complexity has diminished, but, correspondingly, requirements for coordination between ground control

and the aircraft, and between aircraft, are increasing due to the nature of the missions.

That the development of procedures for training in crew interaction is difficult is obvious from the results obtained in the literature. The definition of coordination is imprecise; hence the inability to structure relevant task situations. Moreover, there is ambiguity in response interpretation. For example, more than one equally good solution may apply to a problem. Quite apart from this, a consistent finding from the laboratory researches is that individual proficiency is the most prominent feature of integrated crew performance, and some data suggest that crew coordination results most naturally from the interaction of highly skilled professional airmen during operational missions. Crew proficiency is enhanced by efficient training in defined duty positions. The research data, however, do not usefully identify the variables subsumed under interaction and their influence on task performance. Thus, an important research need is to define the coordination demands that exist in task accomplishment for current and anticipated aircraft/mission combinations.

2. The shifting emphasis toward coordination requirements between aircraft and between air and ground suggests the value of continuing the study of crew communications as indicants of integrated performance. Emphasis should be placed on verbal and nonverbal interaction within a crew and on communications with other units. Such an effort should, for example, yield training procedures to improve comm/nav capability in pilots.

3. Another meaningful research requirement is the development of measures for describing crew interaction and proficient crews. The researches to date have been unable to define scores of crew coordination and reliable criteria of performance. Crew communications measures, for example, have not proved satisfactory.

4. The development of procedures and sequences for training in crew interaction is still a requirement. To this end, a simulation program appears most logical for the research environment because of the difficulties, both technical and administrative, in conducting inflight studies. This strikes a desirable medium between the limitations and restrictions of field research and the sterility of simplified laboratory tasks. The results of controlled simulator studies may be generalized to operations to the extent that the crew task sequences are correct and the experimental conditions carefully chosen. Various issues need systematic exploration. The nature and causes of variation in operational crew performance certainly constitute one issue. Another issue concerns

the effects of critical operational conditions on crew performance, particularly task loading and psychological stress. The relative effectiveness of individual and crew practice, and the distribution of this practice in optimizing performance, are subjects of yet another issue. Theoretically, there is some evidence that crew skills may emerge by individual practice alone. The need is to demonstrate that training in crew skills transfers to the operational task. Allied to the acquisition and transfer of skills is the problem of retention in crew operations. Little information exists on how much forgetting takes place for particular time periods or the rate at which crews forget procedures learned in the simulator.

Visual Aspects in Flying

Enhancing the use of vision in accomplishing flying tasks is the subject for discussion here. The review is limited to two topics: training in visual performance, and training related to the protection of the visual system.

Although vision is the sense modality in flying, not much evidence exists in the literature for enhancing or maximizing the pilot's visual capabilities by training. The limited number of studies in this area make clear that training for the efficient use of vision in flying is difficult and involves a number of complicated problems. While there is some indication that certain aspects of pilot visual performance can be improved by training, it is not at all clear what can be trained or how best to train it. Search and scan techniques are apparently effective only within a given search context, and experience with the specific elements of that situation is required for successful search. Methods for improving target detection skills in the ground environment which emphasize active search lead to improvement in that environment; but there is little expected transfer to operational tasks when motion (in the visual scene) is involved.

Evidence, again very sparse, suggests that training related to the protection of the visual system can be effective in reducing the decrement in performance that is predicted to occur in high intensity light environments such as produced by nuclear blasts.

Training in Visual Performance: Studies appraised here are limited to training for visual search, and target detection and identification. The emphasis is on training for improving visual performance of the unaided normal eye. Since a principal question concerns the extent to which visual performance can be improved by training routines, the discussion is method oriented. No review of studies of the ability of the eye to resolve the detail of objects either on displays or in the environment as a function

of illumination, clutter, contrast, altitude, range, speed, and other variables that affect visual capabilities is attempted. The literature on visual performance as it is affected by these variables under both laboratory and field conditions has already been cataloged by Franklin and Whittenburg (1965).

The essential training problem in visual search is one of training the subject to use his eyes effectively to locate objects in the visual field. This includes a consideration for both the scan pattern used in searching and visual fixation, since the latter determines when target detection can occur. Optimal search strategies for specific situations have been determined analytically and the relative effectiveness of different techniques for search and scan investigated in field studies. In a study involving classified information, Craik (1957, cited in Franklin & Whittenburg, 1965) sought to determine optimal techniques for air-to-sea search operations. His analysis of the data suggested that sweeping the eyes from left to right and right to left along a line 3° below the horizon at about 10° per second should produce the best detection results in this context. Apparently, however, no test of this conclusion was made.

Thomas and Caro (1962) validated analytical search models in the field by testing air-to-ground detection performance with four visual search methods. Subjects, given ground training on how various military targets would appear during flight, were then flown in an L-19 aircraft at 40, 70, and 100 miles per hour at 200 feet over targets placed in an uncluttered area. The method which produced the best detection performance was one in which the subject scanned an area 90° from the line of flight by sweeping his gaze inward toward the aircraft and outward toward the horizon (Side Movement method). Head movement, rather than eye movement, was stressed. The Forward Move, Forward Fix, and Side Fix methods were relatively less effective, in that order, than the Side Movement method. All methods decreased in efficiency as aircraft speed increased. In the study, subjects knew what to look for and how the targets would appear from the air, and these appear to be the critical factors for determining the efficacy of search methods. Other investigations in visual search, summarized by Morris and Horne (1960), lead to the conclusion that optimal search techniques cannot be specified in any general statement. The choice of best technique is a function of the "givens" in a particular search situation, e.g., the target and its characteristics, the context in which it is imbedded, the characteristics of the observer, etc. While, presumably, any training technique which increases the rate and duration of visual fixation per unit of time could be used to improve search performance, the situation is not quite so simple since fixation, while a necessary condition, is not a sufficient one for

search to be effective. Training routines aimed at improving visual search behavior, it appears, must first concentrate on providing the subject with specific experience in actual or simulated situations consonant with visual capabilities, and then concentrate on instruction in scan patterns developed for specific instances. This latter type of "training" may be a matter of simple instructions to the subject.

A series of exploratory studies by the Human Resources Research Office sought to determine if the ability of tank crewmen to detect and identify targets could be improved by training. These studies while not specific to pilot training, contribute to training methods and thus are reviewed here. The first study in the series (Stark, Wolff, & Haggard, 1961) was concerned simply with determining whether training could improve the ability of subjects both to detect and to estimate the ranges of tactical targets in the field. Classroom training on tactical cues on location and distance of targets was followed by field training in which corrective feedback was given. Subjects' performance on a field proficiency test, involving six targets placed at varying ranges, was assessed both before and after training. Training was shown to be effective for improving both detection and range estimation, and the authors concluded that at least some aspects of target detection were trainable.

A subsequent study (Wolff, 1961) compared three different training methods for their effectiveness in teaching target detection to armor personnel. Measures of effectiveness were taken during training and on a transfer task (which was a movie). During training, one group of subjects passively viewed slides of military targets for 30 seconds, presented in an increasing order of difficulty of target detection. In a second group (continuous pointout), the instructor pointed out the location of the target on the slide for the full 30 seconds each slide was presented. A third group (delayed pointout) viewed the slides for 20 seconds before pointout of the target was made by the instructor. Daily tests were administered on each of four training days using slides to assess the efficiency of each training method. At the end of training, all groups were given the criterion task (a one-half hour movie). All three training methods led to significant improvements in target detection performance on the slide tests over the four-day training period. The delayed pointout method was significantly more effective than continuous pointout, and continuous pointout was significantly more effective than no pointout. Differences between methods appeared at the end of the first day of training and continued throughout. There were no significant differences

between methods in the rate of gain in detection skill over days. On the transfer task involving moving targets, none of the methods was superior. In fact, performance of trained groups was essentially no different from performance of subjects who received no training. The author concluded that, at least during training, methods that encourage active searching behavior (delayed pointout) by the subject are more effective for training target detection behavior than methods which only give information.

Subsequent studies in this series investigated the value of other training routines for improving detection performance. Wolff and Van Loo (1962) found that methods of training requiring active subject participation led to an improvement of about 25 percent in detection performance, but also led to an increase in the number of false detections. In the same study, it was found that training transferred to a criterion task (movie) in a very minimal way, as shown by a comparison between trained groups and a control group which simply viewed the movie. Wolff, Burnstein, and Van Loo (1962) sought to extend the training methods previously used with individuals to the training of whole groups of subjects. Results obtained were similar to previous results, in that groups given delayed pointout showed an approximate 30 percent gain in detection performance as a result of training. They also found that a graded progression of training materials (easy to hard) was more effective in improving detection performance than a random difficulty sequence. A final study (Wolff, 1962) found that differential schedules of reinforcement (i.e., giving delayed pointout) did not affect detection performance for a group trained in this way.

Training for Protection of the Visual System: In recent years, considerable interest has focused on protecting pilots from, and preparing them for, the subjective experiences that accompany a tremendous amount of light energy such as that released by detonation of nuclear devices. This interest concerns development of devices to protect the pilot against "flash blindness" and the means for indoctrinating him in the subjective aspects of the phenomenon.

Flash blindness and the attendant "startle" from exposure to the high-energy light release of a nuclear explosion may result in complete loss of mission capability for the pilot. Devices that have been developed for protection against this light energy include light-restrictive

filter goggles, thermal-nuclear cockpit shields, monocular eyepatches, and gold-coated, low-light-transmission visors.[19] Devices have also been developed for simulating light intensities associated with defined levels of nuclear detonation. The rationale of the flash blindness trainer is that it provides the pilot an opportunity to experience these effects, thus enabling him better to cope with the real event when it occurs. One device, developed for the Navy (BioTechnology, Inc., 1966) to form the nucleus of a flash blindness training program, generates a high-intensity electronic flash capable of producing 30 seconds or more of flash blindness. A screen diffuses the blast so that it is more representative of that from a nuclear burst. The screen also allows for projection of a color film of terrain as seen during low-altitude, high-speed flight. A pilot's instrument panel presents tasks representative of those performed in the cockpit and is used to demonstrate the performance decrement which occurs following flash blindness. At the present time, this device is the only flash blindness trainer in operational use. Current Air Force interest in flash blindness training is, however, apparent as evidenced by the initiation of a recent procurement action by the Aerospace Medical Division, Brooks Air Force Base, Texas, for a "Flashblindness Orientation Training Device" to be incorporated into an aircraft simulator.

Research Issues: The efficient use of vision is an important consideration in pilot training, yet very little in the way of data and technique is presently available. The need exists to determine the visual search requirements of pilots in the mission environment and the extent that training can be employed to enhance this visual performance. To this end, a program of research is indicated for developing and validating course material for enhancing visual performance to be used in pilot training schools. Such material is highly suitable for advanced pilot training programs (combat crew training school level) particularly in the Tactical Air Command for reconnaissance, fighter, and special air warfare units.

---

[19] Parker, J. F., Jr., & Bosee, R. A. The success of U.S. Navy equipment development programs in meeting the flash blindness problem. Paper presented at Symposium on "Loss of Vision from High Intensity Light," Aerospace Medical Panel, AGARD, NATO, Paris, France, 16-17 March, 1966.

PERFORMANCE MEASUREMENT

The key issue underlying effective pilot training is the capability for scoring and assessing performance. In essence, the effectiveness of training is dependent upon how well performance is measured and interpreted. This issue is not a recent revelation, for concern for the measurement of inflight performance is almost as old as aviation itself. The National Research Council Committee on Selection and Training of Aircraft Pilots (1942) summarized the early work of aviation psychology as preoccupied with selection, and underscored the failure to conduct research in the air, absence of methods and measures for assessing flying proficiency, and absence of test validation studies. From the end of World War I through the 1930s, the literature reflected this concern for selection (Erickson, 1952a; McFarland, 1942; National Research Council, 1942). A continuing shortcoming of this period was the absence of adequate proficiency or criterion measures. The traditional method of flight performance evaluation relied exclusively on flight instructor ratings for individual maneuvers and on grades of overall performance. Certainly the beginnings of this approach were reasonable. The pilot's job was much simpler and the training programs were leisurely when compared with the era from World War II to the present. Flying was an art requiring no elaborate assessments, only the judgments of expert pilots observing students in the air. Subjective ratings of flying ability continued to be used by the Army, Air Force, and the Navy during World War II as the basis for grading. In the typical evaluation, the instructor recorded his judgment of the trainee's performance on a grade slip after the check ride was completed. The ratings on the instructor-selected work samples and on overall performance were in terms of satisfactory or unsatisfactory or a variant thereof. These grades, together with written comments by the instructor, made up the largest portion of information available on flight performance. Far from being the vogue of "olden days," subjective evaluation is still prevalent today, for example, the Federal Aviation Agency flight checks for Airline Transport Pilot rating (Form FAA-342A) and the Airman Proficiency/Qualification check (FAA Form 3111).

Development of Objective Flight Checks

The subjective flight check lost much of its value as the complexities of flight increased. It resulted in large inconsistencies in the ability to differentiate between student performances. Objections were raised on several important counts. The subjective checks most often failed to produce adequate agreement between independent observers. Check pilots differed in what they considered important to assess, in student expectations (bias errors), in concepts of grading, and in what samples

of behavior should be observed. To minimize these objections, a considerable effort has been expended on making the evaluation system more objective, and a number of studies report the development of objective flight checks. The first systematic program to obtain detailed and objective inflight proficiency measures was begun in 1939 by the Civil Aeronautics Administration Committee on Aviation Psychology of the National Research Council. One contribution of this program was the Ohio State Flight Inventory to evaluate trainee performance in light planes. The initial form presented a series of 5-point rating scales for rating each task in a maneuver. The 1942 version, while still a subjective overall flight check, included some objectively scored items (instrument readings) recorded during flight (Edgerton & Walker, 1945; National Research Council, 1945; Viteles, 1945). Several examples from the 1942 version are shown below.

Throttle:	does ______	or does not ______	keep hand on throttle.		
Planning:	does ______	or does not ______	consider other traffic.		
	drift ______	or no drift ______.			
	corrects ______	or fails to correct ______	drift.		
Levels off:	__________	feet too high.			
Etc.:					

Erickson (1952a), in summarizing the Ohio State Flight Inventory studies, concluded that the project generally gave discouraging results primarily because of a number of formidable problems which included multiple variables difficult to control, small numbers of subjects, and rapid instructor turnover. The research, however, was significant for it provided the first systematic attempt to increase the objectivity of flight proficiency measures. One of the outcomes of this research was the standard flight technique.

The Army Air Force of World War II also employed checklists and grading forms, but these were used as a basis for subjective evaluation. The weaknesses of this technique were well known and attempts were made to minimize the shortcomings. For example, daily grades were used together with check ride ratings and final proficiency evaluation. However, great variation in grading and low reliabilities resulted. The fact that flight

instructors were capable of more consistency in evaluating student performance was, however, demonstrated by Crawford and Daily (1946). They collected the instructors' comments made on the grade slips after each training flight on a sample of over 600 primary flight students. The comments were classified in terms of motor technique, perception, headwork, motivation, and emotional difficulty, and frequencies of comment were tabulated. These scores were correlated with tests from the AAF Classification Battery and with the pass-fail criterion. Greater reliability of evaluation resulted from this than from the use of the grade alone. It was during this period that the Army Air Force Aviation Psychology program began research to develop more objective indices of flight performance, essentially in the form of flight checks involving both objective observations and subjective ratings (Miller, 1947). The work on objective measurement of flying skill (pilot project), benefiting from the insights gained from the CAA program, concentrated on the construction, tryout, and evaluation of a variety of performance measures. Miller (1947) cites 523 different measures that were investigated for contact flying during primary, basic, and advanced phases, and during basic and advanced instrument phases. Several methods of scoring instrument flying were compared. These included the <u>time sample method</u> in which deviations in instrument readout from an established norm were recorded; the <u>range method</u> which represented the difference between high and low instrument readings during a maneuver; and the <u>limits method</u> which recorded the single largest deviation from the correct reading. However, no clear-cut superiority was assigned to any one of the methods. Some effort was devoted to combining separate measures in total scores. In one method, equal weights were assigned in combining the measures; in another, the weight assigned was determined by regression equations to compute beta-weights maximizing the validity of the combined score; in still another, weights which approximated cutoff scores were determined by expert pilot judgments. Comparison of these three methods on maneuvers of levelout and instrument turns yielded essentially similar results. From the total number of measures investigated, the best items were selected and a comprehensive scale of 81 measures of instrument flying skill for the basic level of training was assembled and tried out. However, additional efforts of this kind were halted with the war's end.

The development of objective flight checks carried over into postwar research, resulting in commercial airline pilot proficiency checks, checks for light plane flying, and military programs for evaluating pilot and aircrew proficiency. In 1947, the CAA NRC Committee on Aviation Psychology began a program for determining the airline pilot's job requirements and developing a more standardized method of evaluating flying proficiency. Based on critical incidents, accident reports, and job

analyses, a number of critical factors were assembled and flight checks were built around these components. The checklist forms incorporated the following: Tasks were arranged into a standard flight, uniformly administered; an on-the-spot record was provided of what the trainee did; and critical components of the job were identified. Both objective observation methods involving graphic and pictorial items and subjective items were included in the checks. Objectivity was attained by the use of pictorial diagrammatic aids, quantitative data, and precise descriptions. An example of scoring a maneuver is shown in Figure 6. This extensive program has been reported by Gordon (1947; 1949) and Nagay (1949; 1950).

A well-conceived program for improving methods of grading pilot performance was that accomplished by the Basic Pilot Training Research Laboratory of the Human Resources Research Center (Smith & Flexman, 1952; Smith, Flexman, & Houston, 1952). An objective technique was developed for recording student pilot performance. One outcome of this research was the Performance Record Sheet (PRS). The PRS was designed to describe performance as completely and objectively as possible and was not meant to evaluate that performance. Its construction was based on an analysis of the maneuvers in the Air Force Primary Pilot Syllabus (T6-G airplane) and on interviews conducted with experienced flight instructors. From this analysis, items descriptive of the critical elements of performance in the maneuver were constructed. For example, the "steep turn" maneuver was broken down into the following elements:

<u>Entry</u>:

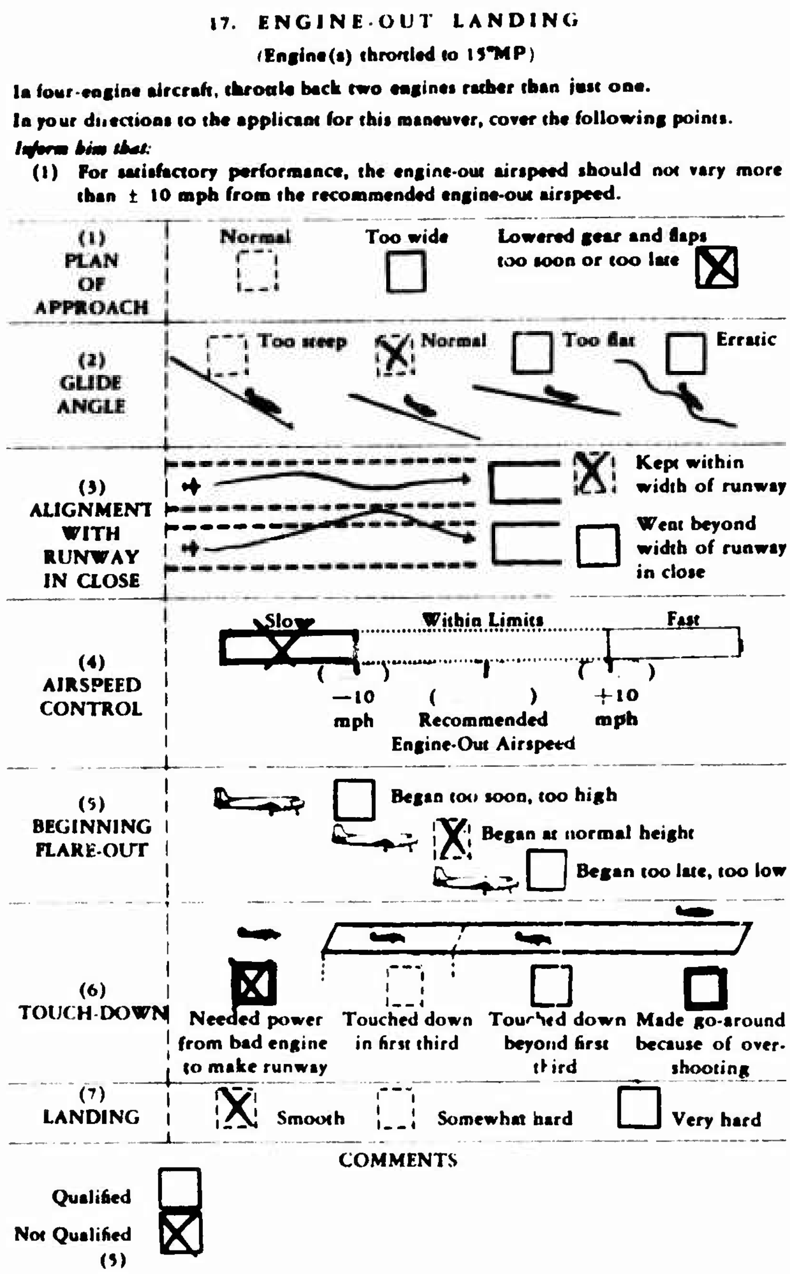
Looks ______
Bank (60° minimum) ______
Altitude held ______
Coordination ______

<u>Maintaining</u>:

Bank (60°) ______
Altitude control ______
Looks around ______
Coordination ______
No high speed stall ______

<u>Recovery</u>:

Direction ______
Wings level ______
Coordination ______
Altitude held (until airspeed 140 mph) ______



17. ENGINE-OUT LANDING

(Engine(s) throttled to 15"MP)

In four-engine aircraft, throttle back two engines rather than just one.

In your directions to the applicant for this maneuver, cover the following points.

Inform him that:

(1) For satisfactory performance, the engine-out airspeed should not vary more than ± 10 mph from the recommended engine-out airspeed.

(1) PLAN OF APPROACH — Normal; Too wide; Lowered gear and flaps too soon or too late

(2) GLIDE ANGLE — Too steep; Normal; Too flat; Erratic

(3) ALIGNMENT WITH RUNWAY IN CLOSE — Kept within width of runway; Went beyond width of runway in close

(4) AIRSPEED CONTROL — Slow; Within Limits; Fast; −10 mph; Recommended Engine-Out Airspeed; +10 mph

(5) BEGINNING FLARE-OUT — Began too soon, too high; Began at normal height; Began too late, too low

(6) TOUCH-DOWN — Needed power from bad engine to make runway; Touched down in first third; Touched down beyond first third; Made go-around because of over-shooting

(7) LANDING — Smooth; Somewhat hard; Very hard

COMMENTS

Qualified

Not Qualified

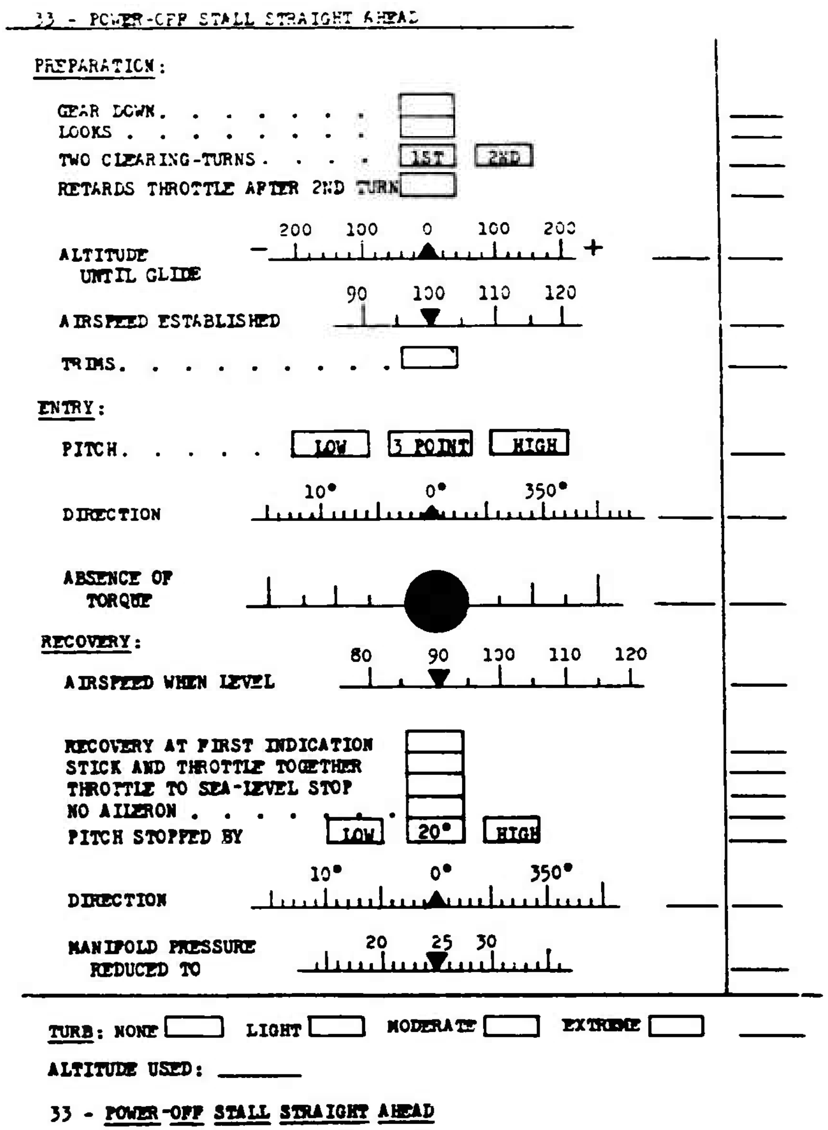
(5)

Figure 6. Example of a Maneuver Scored in the Airline Pilot Flight Check (from CAA Flight Check Manual, Department of Commerce, Washington, D.C., April 1950).

The items selected were those determined to be most important and reliably recorded. Two types of items, in about equal proportions, made up the PRS: scale items (in or out of a predetermined tolerance), and categorical items (did or did not accomplish the item). A sample of a Performance Record Sheet showing a maneuver (from a total of 78 identified) and item definition is provided in Figure 7. The development of the technique, however, did not progress beyond the stage of its use as a research tool. The authors felt that the Performance Record Sheet could provide normative and reliability data for use in constructing an adequate proficiency measure for evaluating student performance. Also, information collected with the PRS could be used to set realistic standards for student training, as later demonstrated (Houston, Smith, & Flexman, 1954).

The Navy began in 1951 to develop and evaluate objective inflight grading methods for two stages of naval air training (Wilcoxon, Johnson, & Golan, 1952). The approach, patterned after Gordon's check (1949), involved standardized check flights, clear definition of maneuvers to be performed, itemized objective record of the trainee's performance, and inflight (or soon thereafter) marking of performance. Also, instructor comments were collected. Check flight grades were collected on two identical flights for each trainee, administered by two different flight instructors. However, the check proved no more reliable than the traditional Navy flight check. Low reliability was attributed to daily student variability rather than to measurement error. Danneskiold and Johnson (1954) substituted seasoned naval aviators undergoing training in the instructor basic training unit as subjects, to check the conclusion that day-to-day fluctuation in student pilots was responsible for the low reliability. The assumption was that experienced aviators were less variable in day-to-day performance. Although the ride-ride reliability improved, the study supported the hypothesis that performance of flight skills varies considerably between flights.

During 1957, Army Aviation developed a standardized, relatively objective check flight for primary and basic light helicopters to replace the traditional flight check. This work, conducted by the Human Resources Research Office (Duffy & Colgan, 1963; Greer, Smith, & Hatfield, 1962), was initially guided by the flight check developed by Smith, Flexman, and Houston (1952). Identification of the most frequent student errors was based on an analysis of grade books and interviews with instructors. Also, intermediate and advanced flight maneuvers were analyzed in flight to determine fundamental components of each maneuver. From these data, judgmental and descriptive scales on each performance component were then developed. These scales, called Pilot Performance Description Records (PPDR) provide a standardized record of student performance.



33 – POWER-OFF STALL STRAIGHT AHEAD

PREPARATION:

Gear Down: Should be considered correct if student lowers gear before starting maneuver.

Looks: Def. 10. An error will be scored if student fails to look in either turn.

Two Clearing-Turns: Def. 11.

Retards Throttle after 2nd Turn: Throttle should be fully closed prior to completion of 2nd clearing-turn.

Altitude Until Glide: Def. 5. Variations should be scored from altitude where power is reduced until glide is established.

Airspeed Established: Attitude should be established which will result in airspeed of 100 MPH with throttle fully retarded. Airspeed should be scored immediately before changing pitch to stall position.

Trims: Def. 9. Should be considered correct if completed any time prior to stall.

ENTRY:

Pitch: Def. 12. Aircraft should be put into slightly nose-high attitude generally corresponding to attitude in round-out for 3-point landing. Attitude should be scored in appropriate box.

Direction: Def. 6. Should be scored from completion of clearing-turns until stall occurs.

Absence of Torque: Def. 1. Should be scored from completion of clearing-turns until stall occurs.

RECOVERY:

Airspeed When Level: Airspeed is graded immediately upon recovery to level-flight attitude following stall. Actual airspeed is scored.

Figure 7. Sample of the Performance Record Sheet (PRS) (from Smith, Flexman, & Houston, 1952).

For example, the primary check PPDR consists of 17 flight maneuvers with a total of 236 separate items of flight performance. Each item is scored separately and then the maneuver is graded on a 4-point scale (above average, average, below average, unsatisfactory). The objective scales measure easily defined aspects of performance (instrument readings, position to the ground); the subjective scales measure aspects that must be judged by an observer (approach angle, area selection). A sample page from a PPDR is shown in Figure 8. While the PPDR does not provide a complete description of performance, it encompasses a considerable range of information to be recorded during a single flight period.

The Air Force currently employs a formal program for evaluating aircrew proficiency, called the standardization/evaluation check flight (stan/eval check). The stan/eval is made up of a ground phase which involves written examinations, and usually a check ride in the simulator for the aircraft; and a flight phase which involves inflight evaluation of elements of the overall mission profile. This is usually followed by a critique where the evaluator discusses the good features and the discrepancies in performance with the crew being examined. The stan/eval check is complex. Written examinations cover a substantial range of knowledge in a specialty and involve both open- and closed-book testing. Simulator checks concentrate heavily on the pilot's ability to handle simulated emergencies, some of which cannot be assessed safely in the air. Instrument checks are also included. The prime emphasis in the stan/eval check is on demonstrated crew performance in the air. The standards of performance are established in the 60-series manuals for the air command (51-series in the Strategic Air Command), with a volume for each aircraft type. The criteria include detailed performance descriptions of each aspect evaluated, limits of error permitted, and an adjectival rating for each level of performance. For example, the SAC stan/eval check (SACM 51-4, 1966) requires the check pilot to rate on four levels of performance: highly qualified (H), qualified (Q), conditionally qualified (C), or unqualified (U). Performance is categorized into areas either by phase of flight or aspect of mission accomplishment. In most instances, these areas are further divided into subareas for more definite analysis of performance. Each item of behavior within a skill area is scored on a worksheet in terms of objective references (e.g., tolerances, percentages, numerical ratios, or adjectival descriptions). For example, in the instrument departure phase, evaluating the B-52 pilot's ability to achieve climb and leveloff airspeed, leveloff altitude, course holding, and TACAN arc, the scores for highly qualified (H) require performance of: ± 5 knots or ±.01 mach airspeed; ± 100 feet of altitude; ± 3 degrees in course; and TACAN arc of 1. Conditionally qualified (C) requires ± 15 knots or .03 mach airspeed; ± 200 feet of

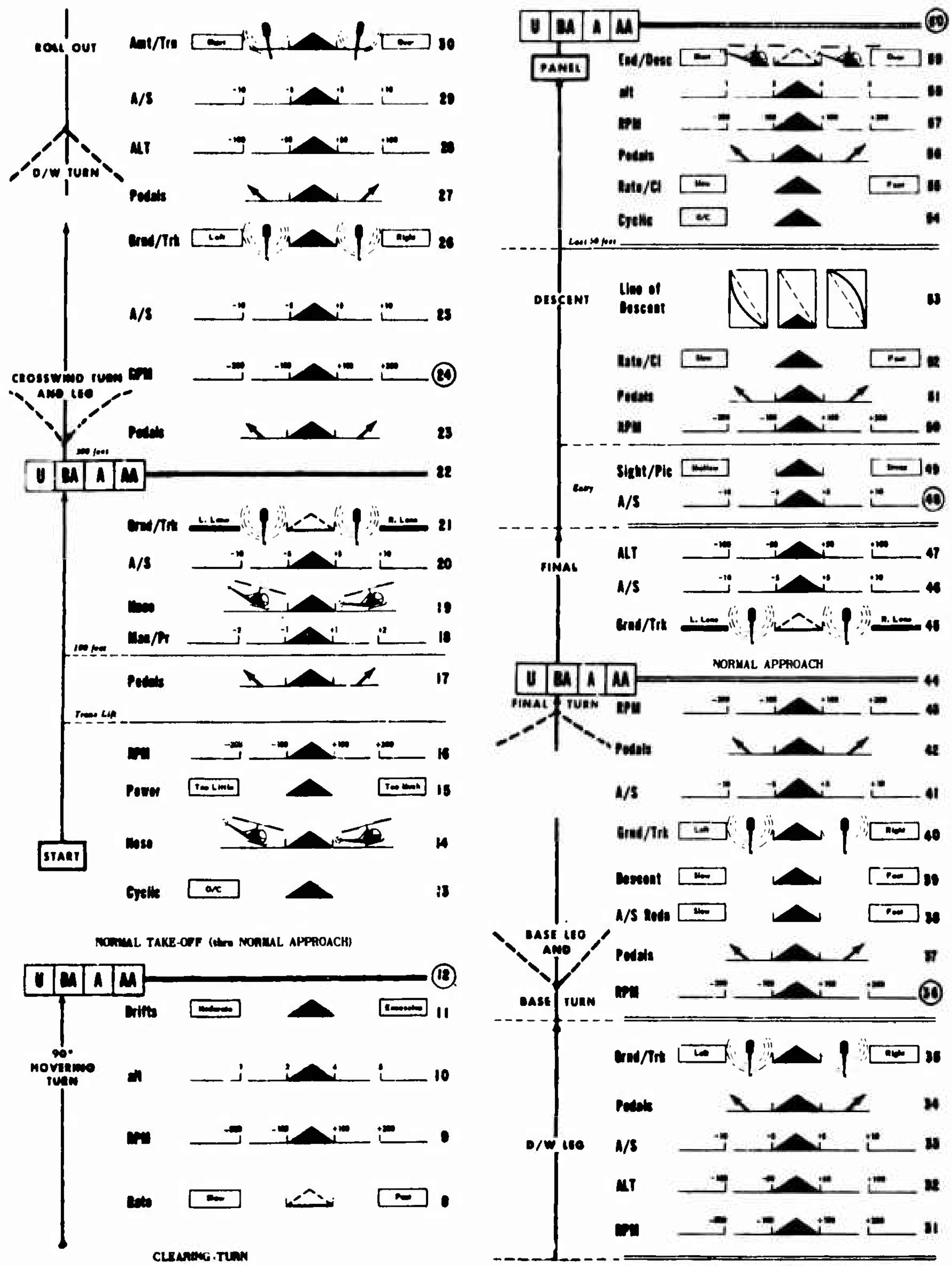


Figure 8. Sample page from the Pilot Performance Description Record (from Duffy & Colgan, 1963).

altitude; ± 9 degrees in course; and TACAN arc of 3. The evaluator determines the score for each item of performance on the basis of these criteria. Those behavior items and skill areas in which inadequate performance will seriously compromise the success of the mission are designated as critical. An unqualfied score on any critical item yields an area score of unqualified, which results in overall unqualified score for the standboard evaluation.

The inflight evaluation format is an objective record based on subjective judgments. The record of crew performance is kept on a prepared worksheet which is a structured objective checklist on which the evaluator makes specific observations about performance and makes detailed notes on the behaviors observed and discrepancies encountered during the evaluation. So far as an overall statement can be made, the stan/eval program is practical (although costly in time and money) in answering the basic question of whether or not this crew is operationally/combat ready. The utility of the program is evidenced in the reduction in accidents and the pointing up of soft spots where additional training is required.

As a system for proficiency evaluation, though, the reliability and validity of the stan/eval are not known. Few comparisons of reliability have been made. The features of subjectivity, evaluator biases, changing criteria (which are, in part, based on logical and practical thinking modified by experience, and in part, on modifications in aircraft), and the probable contingencies that arise during flight to prevent equivalence in testing are classic reducers of reliability. The question of validity defies description because of the nature of the combat mission and because crews are graded on the basis of constantly changing, more immediate criteria of success. It appears doubtful that the opportunity to evaluate performance in the criterion situation (i.e., fulfilling the tactical mission) will present itself except in the grossest sense.

<u>Reliability of Flight Check Systems</u>: Desirable advances toward objectivity in measuring performance have been made in the various flight checks described. Although both subjective and objective measures are obtained, the subjective judgments are made on relatively smaller, well-defined aspects of performance (e.g., application of power during a maneuver, approach angle, etc.) than was the case in the older traditional flight checks. This has resulted in improvements in the observer-observer reliability of the check flights.

In the Army Air Force research (Miller, 1947), correlations between grades on the traditional subjective check flight given at the completion of training and those given earlier in training were poor, with

coefficients often in the twenties. Similarly, in the traditional Army flight check system, the average training grade-check grade relationship was about zero. The correlations were .35, .08, and .09 for the presolo, intermediate, and advanced stages respectively (data from 1956 and 1957, cited in Greer, Smith & Hatfield, 1962).

Unfortunately, the reliability of the more objective checks has been less than hoped for, being generally low and/or showing large fluctuations. Of them all, the airline pilot proficiency checks demonstrated the greatest reliability. On the revised experimental test, using a sample of 26 pilots, the ride-ride reliability was .76 with an observer-observer reliability of .86 (Gordon, 1949). It should be added, though, that these high correlations were obtained from data on experienced pilots in the more procedurally oriented commercial flight environment. The ride-ride relationship in the Air Force checks (Smith, Flexman, & Houston, 1952) ranged from .17 to .67 with a .50 average. In a Navy study (Wilcoxon, Woodbury, & Golan, 1952), a ride-ride relationship of .33 in the instrument stage and .31 for the primary syllabus of training was reported. The low correlations were attributed by the authors to variability in student performance from one ride to the next. In fact, the evidence from the various studies cited suggests that the pronounced variability in the checks is due less to errors of measurement than it is to students' ride-to-ride or day-to-day fluctuations, to changing wind and weather conditions, and to differences in airplanes.

The conclusions to be derived from this significant body of inflight research are that the scoring methods of the subjective flight checks and the more objective flight checks yield low reliabilities and a less than desired capability for differentiating between students. The subjective checks suffer additionally from greater error in grading, since individual check pilots make their evaluations in terms of their own standards, and the variations among these pilots are substantial.

Although gains have been made in developing more reliable, more comprehensive, and more diagnostic flight checks, certain features conspire to reduce the capability for measuring and assessing inflight performance. These are summarized as follows:

<u>Check pilot biases</u>. Evaluation is wholly based on the judgment of the examiner, and various biases at one time or another influence the results.

<u>Flight environment</u>. It is difficult to measure and evaluate performance in the air. Pilot performance is affected by a variety of interactions involving contingencies in flight and changes in individual

reactivity (intra- and interday fluctuations in trainee performance), to which may be added hazard and safety features as well as interpersonal aspects between the examiner and the trainee.

Precise measures. An adequate number of effective measures for describing performance is not available.

Validity of the checks. The validity of a proficiency test is due in large part to the accuracy with which the job has been analyzed and to the selection of the critical events to be measured. No indication of validity of the flight checks was discernible in the studies cited. Nor can validity be easily expressed. At present, pilot training research has been unable to define precisely the pilot's job and, hence, unable to specify the critical behaviors to be assessed. Validity, although indeterminate, is assumed to be adequate based on subject matter expertise about flying.

Summary of Inflight Measurement: The research on flight check development has shown a consistent trend toward increasing objectivity in scoring performance. Yet, with perhaps the exception of the research accomplished for Army Aviation (helicopter flight checks) and a conglomorate of inputs to the Air Force Standardization/Evaluation program, none of the evaluation instruments is in use today. The obvious question is: "What are the reasons for not using these research results?" There are several. The systematic flight checks require special training of the instructors. Also, flight instructors resist these techniques because they require more "head in the cockpit" time than they are willing to allot. Finally, there is a certain natural resentment against the regimentation of setting up and observing this event at this time. Flight instructors intuitively feel they know best how to assess training progress and outcome.

In a more general vein, the evaluation of inflight performance is a long way from being effectively achieved, and less than complete information is provided by present measures and methods. The prevailing case is that decisions on what aspects of behavior to sample, and when and under what conditions observations are made, are left to expert judgment. Measures obtained are often indeterminantly associated with overall proficiency. In many instances, measurement is sufficiently difficult that the practice is to obtain what is measurable rather than what is desired. Another serious difficulty with flight measurement is the frequent inability to detect and assess differences in performances when they, in fact, exist. This results in haphazard conclusions. Certainly a good deal of the present ambiguity in flight research can be attributed to this feature. The overwhelming problem continues to be the inability to structure the inflight environment so that accuracy, reliability, and validity of measurement are within tolerances. This has been the impetus for developing a full-scale measurement capability in simulators with provision for automatic scoring.

Quality Control

Procedures have been installed in aviation training for establishing quality control in the production of pilots. Measurement data provide the basis for determining needs for changes required in order to maintain or adjust the quality level of the pilot product. A systems approach to training is employed with emphasis placed on precise statements of training objectives and a capability for precise evaluation of the program. The criteria for evaluating such a program are based on objectives defined in terms of pilot job requirements, i. e., what the trainee should be able to accomplish and to what standards of proficiency. Achieving such absolute standards of performance requires criterion-referenced measures which describe the degree of competence attained by the trainee independent of the performance of others, i. e., comparison of the individual with the capabilities of the training system (see Eckstrand, 1964; Glaser & Klaus, 1962).

Publications are available which define quality control practices in naval aviation (Berkshire, 1965; Shoenberger, Wherry, & Berkshire, 1963) and in Army aviation (Duffy & Colgan, 1963; R. Smith, 1965). Smith has described some general guidelines used for quality control in Army aviation training programs. The elements required in a successful program are: (1) statements of training objectives based on job requirements, (2) accurate and appropriate proficiency measures, (3) effective communications concerning student performance, (4) procedures for corrective action, and (5) supervisory support. Duffy and Colgan (1963), continuing from the work of Greer, Smith, and Hatfield (1962), report on a system of quality control for Army helicopter training. Check pilots, not on the instruction staff, are assigned students at random for flight check in which the Pilot Performance Description Record is used. Each scored PPDR is machine processed and a percentage of error is computed for each maneuver. These data are compared with a standard of performance which is the average performance of a number of recent classes. Thus, individual performance can be diagnostically compared with school standards. Also, a measure of instructor effectiveness is afforded by observing trends in the performance of the individual instructor's students. Logically, systems such as this provide information not only for passing or failing students, but also for determining if the current class is up to defined school standards and for determining instructor proficiency. Just as logically, such systems require precise statements of training goals, uniformity in instructional technique, reliable and valid measures of performance, and effective assessment practices. These capabilities are not optimally achieved with current training technology.

Assessing Aptitude for Military Flying: A group of studies has attempted to identify aptitude for military flying by investigating the effects of light plane flying or indoctrination flights upon subsequent student performance. Although conducted for selection purposes, i.e., prediction of later success as a pilot, the studies are of interest here in that they provide information on rate of progress during initial flight training. (Problems of pilot selection and pilot performance measurement have a common base in that the methods used imply measurable requirements and characteristics of the successful pilot.)

As part of a program to improve the initial selection and screening of student pilots in the Air Force, a study was begun in 1951 (Ericksen, 1952b) to develop a light plane proficiency check for the purpose of identifying students with aptitude for military flying. An objective checklist of over 400 items was developed based on interviews with pilots, analysis of grade folders, and a survey of the primary phase flight syllabus. The method involved a checking of the errors made in each maneuver which yielded an overall "points-off" score. A preliminary reliability study on 28 private pilots in which check pilots administered both rides on different days yielded a ride-ride reliability of .31. Unfortunately, the results are spurious, since two different types of light plane were used and the check pilots were not thoroughly familiar with the grading form. Also, the weather during the test was extremely cold. No further information could be found on the use of this flight check for selection purposes or on the predictive value of this technique for later military flying success.

Two studies (Cox & Mullins, 1959; Mullins & Cox, 1960) summarized the 1956-1957 Air Force ROTC Flight Indoctrination Program (FIP). This program (given by civilian flying school operators approved by the Federal Aviation Agency) provided 36.5 hours of flight training (20 hours dual, 16.5 hours solo) in aircraft rated at less than 200 horsepower, and 85 hours of ground school. The issue was whether or not participation in FIP improved a man's chances for successfully completing flight training in the Air Force. FIP and non-FIP officers were compared as to proportions eliminated from primary pilot training. From a total of 357 FIP officers, 11 percent were eliminated: 5 percent for flying deficiency and 6 percent for motivational reasons. For 209 non-FIP officers in primary training, the percentages were 26, 19, and 7 respectively. The results suggest that FIP training reduced the proportion eliminated in primary flying due to flying deficiency, but did not affect interest in flying. In basic flying training, however, the differences between FIP and non-FIP did not hold up. The suggestion is that FIP has value as a selection device, but the evidence does not answer the question of the transfer of training value of this to later aircraft in the training program (see page 56 of this report).

In a study by Berkshire and Ambler (1963), 196 Navy students received a one-week flight indoctrination course prior to the formal preflight school. Four flights totaling 5.9 hours were given plus ground school. Each trainee in this experimental group was matched with a control based on AQT scores. The results indicated that indoctrination flights reduced attrition. The authors added, however, that the "special treatment" aspects could not be ruled out to account for the enhancement effect. The instructor's observations on the trainee abilities for flying had high validity in predicting later failure. For example, of 32 trainees designated as poor risks (airsick on all four hops, highly anxious and tense, or with obvious flying inability), 24 did not complete the flight program. Although the data are crude (subjective opinion on a 7-item questionnaire), the technique may serve as an additional screener of potential attrition cases.

Scoring Capabilities in Simulators

The shortcomings of inflight measurement of proficiency, with no obvious improvements in sight, have convinced all but a few that the only hope for effective measurement is the synthetic ground environment. Development of a measurement system for simulators is the trend for the immediate future. Unfortunately, few studies are reported in the literature which bear directly on implementing an objective measurement system in simulators for assessing pilot performance. As indicated earlier, proficiency measurement in aviation has relied heavily on the human observer for inflight assessment. The same holds true for the simulator except that the simulator is not used in any substantial way for measuring aircrew performance. Its prime use is in teaching normal and emergency procedures for the aircraft represented, and instructor critique is the means of supplying performance information to the trainee. Nor are simulators in use today equipped with a well-defined scoring system. At most, recording devices are available which are used as adjuncts to assessment, e.g., providing knowledge of terminal performance or segment of flight to the trainee. The flight path recorder and the approach recorder, which show a plot of ground track and the path to a radio station, respectively, are two examples of this type of equipment. Scoring and recording equipment have been treated as research items belonging in the laboratory and not in the operational environment. Consequently, few research studies deal with objective scoring in the simulator.[20] Those that

[20] Part-task trainers are not considered in this discussion since they are not well-suited for proficiency measurement which approximates real-world events. Performing on a part-task trainer is easier than the corresponding flying job since it presents only a portion of what the pilot

do, have focused on objective recording of portions of tasks in specific skill areas of the pilot's job, with the recorded data analyzed at a later time. Representative studies of these initial attempts at objective scoring of certain flight aspects in the simulator are described below.

An early study analyzing the problems of scoring trainee performance in simulators was conducted for the Navy by Danneskiold (1955). The purpose was to improve operational flight trainer grading procedures for use by instructors, and part of the effort was to determine the feasibility of mechanical scoring methods. A number of mechanical devices for scoring were investigated which provided either graphic records or a count of errors from some preset standard. These were: the Link Counter SM1060 (continuous graphic index plus deviation counter), Pennsylvania Control Movement Recorder (movement of controls), work adders to score frequency of movement, devices for measuring pressure exerted, motion photography, light panels (indications of sequence out of order), and flight path recorders ("crab"). Pertinent mechanical measurements identified included: heading, angle of bank, altitude, pitch, airspeed, control movements (frequency, distance traversed, amount of pressure), and procedures (sequential correctness, completeness). The conclusion reached at that time was that few mechanical scoring methods existed which offered practical advantage for simulator usage. Although accuracy of measurement was a value of these devices, their limitations precluded effective employment except in precisely defined instances. Danneskiold felt that the mechanical techniques suffered from various difficulties and enumerated the following: (1) inflexibility--mechanical devices provide a record of only a few specific indices of performance, being unable, for example, to record sequential control movements and smoothness in control; (2) scoring devices are too cumbersome; and (3) the scores do not reflect larger, more meaningful aspects of behavior underlying flying skill, since they are applicable only to aspects of flight which are directly linked to the simulator computer. Thus, improvement in scoring using mechanical devices is best achieved by concentrating on task aspects most predictive of overall performance, and on those task aspects which human observers have difficulty recording accurately (i.e., measurement of performance aspects low in reliability but adequate

---

does. Consequently, part-task trainers yield spuriously high measures of proficiency. They should, of course, possess scoring equipment, but such equipment is used primarily for promoting learning, e.g., knowledge of results about performance. (See, for example, Adams & McAbee, 1961.)

in validity). Since that time, remarkable advances have been made in hardware design which, for the most part, have obviated Danneskiold's objections to mechanical scoring.

Some experimentation (Swanson, 1957) has been accomplished using oscillograph recordings of selected indices of pilot performance as aids to evaluation. Swanson presented a report on the feasibility of using a six-channel oscillograph for recording pilot performance in the B-52 simulator. A limited number of normal and emergency activities in piloting skill (i. e., nonprocedural activities) appeared amenable to accurate representation by recording equipment. Webber (1958) proposed that flight instrument data could be ideally recorded and processed for use in assessing performance. He suggested a system that would digitize flight instrument values and record them in a form that could be processed by a high-speed digital computer (i. e., rapid reduction of vast amounts of data).

The engineering feasibility of automated scoring in the simulator has been heightened with the advances made in the design of digital computers used in simulators, and aircraft builders are seeking such a capability in their pilot performance research studies. Sikorsky Aircraft Corporation, for example, has such a requirement for its variable stability helicopter research simulator (Smode, Vallerie, & Kelley, 1963). Curtiss-Wright Corporation (Benenati, Hull, Korobow, & Nienaltowski, 1962) has developed a design for an automatic monitoring system for the flight simulator utilizing available engineering means for recording pertinent mission parameters. Trainee performance can be scored on the basis of comparing the parameters monitored with the programmed performance standards, e.g., errors in performance are printed out for use by an instructor. The report, however, is limited to a specification of principles involved in monitoring and scoring selected parameters and does not deal with the design of a proficiency measurement system.

Preliminary work on automatic scoring of flight task performance was conducted with the Universal Digital Operational Flight Trainer (UDOFT) (cited in Aerospace Medical Research Laboratories Report P-40, 1963, not identified by author but prepared by Robert Buckhout and Theodore Cotterman). Several jet pilots flew the simulator in the F-100A configuration from takeoff to altitude and maintained a holding pattern. Pilots were scored on the ability to abort on takeoff and on airstarts following simulated flameouts. The study demonstrated the feasibility of automated scoring in digital simulators.

A recent study by Bowen, Bishop, Promisel, and Robins (1966) is significant because of its attempt to automate the scoring of portions of the pilot's tasks in Navy training. The research investigated the effects of two training treatments on pilot performance in a Navy Operational Flight Trainer (OFT) simulating the A-4 aircraft. As part of the study objective, scoring devices and procedures were devised which would provide a reliable basis for assessing pilot proficiency. Two groups of ten pilots each, transitioning to the A-4 aircraft, were given three simulator training sessions spread over a period of 20 weeks. The control group of ten pilots flew conventional flights which included normal feedback from the instructor (dialogue between instructor and student). No set pattern existed for this dialogue, which was loose in structure, mostly qualitative in content, and not recorded. Scoring was either satisfactory or unsatisfactory for each procedure. For the same flights, the experimental group received several levels of augmented feedback based on the scoring procedures developed. Knowledge of performance information was given bit by bit (error was identified when made), by numerical grade for each emergency procedure (series of tasks), and by an overall summary score at flight completion. The results indicated that objective scoring information given immediately to the pilot in a useful form, heightened performance. The finding fits well with the literature on the effects of augmented feedback on task performance (see, for example, Annett, 1961; Smode, 1958). The conclusion is best regarded as suggestive, however, since the study was conducted in the operational environment on a "noninterference" basis; hence, problems in strict experimental control were experienced. Matching of subjects in the two groups was not possible. Also, the three OFT sessions were spread over a period of 20 weeks, and the authors contend that so much else was going on in the training course that the OFT experience tended to be "minimized." An enhancement effect for the experimental group cannot be discounted either, since these pilots may have been "urged" to know the emergency procedures well. Nonetheless, this study is significant for its attempt to implement an objective automated scoring system (albeit rudimentary) for the simulator. The scoring instrumentation which was designed and added to the OFT was as follows: (1)A panel of lights which displayed the sequence of events performed during emergency procedures. A buzzer sounded when a sequence was in error. The displayed sequence of events was manually recorded on a scoring sheet. (2) Electric stop clocks programmed to give response times and total time in completing a procedure. (3) A Brush eight-channel pen recorder which scored and documented discrete, continuous, and mixed events. Comments made by the trainee or the instructor were also recorded.

Three sets of independent scores of pilot skill in the simulator were developed.

Emergency Procedure Scores:

S (sequence score)	Number of steps accomplished in a procedural task, weighted for number and difficulty of steps completed and not completed.
P (proportion score)	Proportion of steps correctly accomplished to the total number of steps required (simplified version of the S score).
B (binary score)	Proportion of the number of total procedures correctly accomplished.

Time Scores:

RT (response time)	Time to complete correctly the first step in a procedure.
TT (total time)	Time to complete an entire procedure.

Aircraft Handling Scores:

C (control score)	Rating from polygraph records of control effectiveness over aircraft and engine parameters.
Instructor Ratings	Proportion of satisfactory procedures.
Self (pilot) Ratings	Five-point scale for rating emergency procedures, precision flight, and general pilot ability.

The authors suggest that valid simulator measures require that the trainee be exposed to a multiplicity of tasks and events similar to real flight conditions and that the difficulty level be equivalent to the more difficult aspects of real flight. In this way, the pilot will perform in a pattern of priority as in actual flight (e.g., timesharing attention

shifts, anticipation of events, etc.). The pilot would be practicing in a realistic way the actual skill requirements of flight.

Devices are recommended for objectively recording the following data:

Procedural sequences and computation of a score which accounts for the relative difficulty of sequential steps.

Deviations from required flight and engine parameters. Such data may be inspected for radical deviations or used in a computational program as part of a score for manual control.

Response time to unexpected situations.

Accuracy of precision flight, including a composite score.

Accuracy of navigation.

Finally, the scoring devices should possess a high-speed capability with hard-copy printout in near-real time. A computer-based system is preferred because of the need for updating and improving the various scoring and recording programs.

The achievement of an effective measurement capability is contingent upon having accurate and reliable scores describing overall ability as well as task and task element performance. Such versatility in scores permits diagnostic assessment of those strengths and weaknesses in performance which contribute to the total score. This implies the advantageous use of objective and automated scoring and recording techniques. By all logic, this makes the simulator the prime instrument for evaluating flight performance. Several studies have made the point that it is high time a comprehensive scoring capability was developed in the simulator, enunciating a rationale which emphasizes the payoff issue of precise scoring of all critical aspects of flight heavily involving objective, automated scoring, since only minimal restrictions are placed on power, weight, and space. In essence, these studies, which take into account the somewhat contradictory functions of the simulator for training and for performance assessment (see Gagne, 1962), present systematic schemes for achieving a fully developed measurement capability in the simulator. These studies are described below.

Aerospace Medical Research Laboratories Report P-40 (1963) describes in detail the increasing need for automatic scoring equipment

for assessing the proficiency of aircrew in flight simulators and provides recommendations for developing such a scoring capability. A number of steps are discussed in constructing scores for proficiency evaluation. The steps which follow summarize the requirements for the score development program stipulated in the report. Beginning with an understanding of the purpose and the use of the scores, defining, and classifying the behaviors to be measured is an initial undertaking. Means for quantifying these behavioral elements must be established. Thus, the pilot's job must first be partitioned into manageable units of behavior (tasks, task elements, etc.) which are observable and (in most instances) quantifiable. Criteria or standards of performance must then be developed since quantifying behavior involves examination of mission requirements and selection of those parameters which are useful in evaluation and which can be efficiently obtained. Following this, the scores must be tested to see how well they predict performance, i.e., was the correct selection of scores made? Finally, how well the scores stand up under repeated use must be determined. This systematic development is a big order, since methodologies are lacking to accomplish each of the requirements completely. The study continues by citing the need for systematic inquiry into the behavioral and engineering problems generated by a complete, highly automated simulator scoring capability. For example, the problems of automatically scoring the critical, meaningful aspects of complex tasks have not yet been thoroughly analyzed. The research needs include selecting or developing recording procedures, selecting measures, developing testing programs, determining the applicability of measurement operations to task structure, and predicting mission performance. It is not simply a problem of identifying and using available scoring equipments, for great damage can be wrought by a poorly conceived measurement system in terms of user faith and just pure error.

Various proposals have been made for developing complete measurement-assessment packages, integral to the synthetic ground environment, as the means for overcoming present lacks in measurement capability. These proposals emphasize logical and systematic development, beginning with the precise determination of measurement objectives and culminating in an integrated behavior/measures/scoring hardware array for evaluating performance. Ultimately, the automatic scoring provision in simulators is envisaged to measure quantity as well as quality of performance in order to provide the necessary basis for good instruction. The proposal by Smode, Gruber, and Ely (1962; 1963) is cited here as an example of those studies that have the common goal of specifying the need and requirements for a complete performance measurement capability in simulators, largely automated. These writers propose a sequence of logical steps for implementing a measurement system. The initial

effort is the specification of the criteria for what constitutes proficient performance. To a major extent, these performance standards define the tasks and elements to be measured, indicate the terms for expressing measures, and provide a base against which measurement data can be compared in evaluating proficiency. Specifying performance standards requires knowledge of the system, its objectives, missions, and the jobs and tasks which comprise it. A detailed account of the procedures for setting up a measurement capability is presented. The essence of these logical steps is outlined briefly below.

Conduct a system and job analysis. A thorough knowledge of the tasks and behavioral requirements within a system context is a necessary first step in the design of a measurement system.

Identify important and critical tasks. The evaluation of human performance requires knowledge of what the significant behaviors are and their characteristics relative to measurement. Further, decisions must be made whether to select more discrete units of behavior or more comprehensive segments of performance. The desired product is a list of things worthy of measurement. A behavioral classification for the purposes of measurement is a difficult undertaking and no clear and firm guidelines exist. A number of studies have developed behavioral taxonomies for the purpose of relating task classes to principles of training (see, for example, Fitts, 1962; Miller, 1962; Smode, Gruber, & Ely, 1962; Willis, 1961).

Determine performance requirements for the important tasks. The consequences of performance must be well understood and performance requirements stipulated prior to the selection of measurement classes.

Select measures appropriate to the behavior to be evaluated. A variety of measures are available for describing specific behavioral sequences in performance. It is important to determine the types and number of appropriate measures required and the manner in which they will be obtained. The selection of measures pertinent to the purpose should include both diagnostic indicants of skill areas and an overall measure of performance.

Determine conditions under which to measure critical tasks. Both environmental and task conditions must be selected. These conditions should be representative of the range found within anticipated operational situations.

Decide on techniques for recording measurement data and for combining separate measures. Means (hardware versus observer) for obtaining appropriate measures and means for displaying and recording

the data must be determined. Also, ways must be established for combining measures whenever appropriate.

Since different kinds of measures are not equally applicable to different kinds of behavior, decisions must be reached on those measures required for critical behaviors. To this end, a matrix is suitable, showing the interactions between behaviors and measures. Table II shows the form of such a behavior-measurement matrix for determining those measures most applicable to each type of activity.

TABLE II

APPLICABILITY OF MEASUREMENT OPERATIONS
TO JOB BEHAVIOR

Behaviors to be Measured		Types of Measures
	Elemental Tasks	
	Complex Tasks	
	Job Segments	
	Overall Mission Behavior	

With a knowledge of the given system and the criteria for system performance, the possible measures that will describe a known behavior can be determined.

Instrumentation is a key factor in an automated measurement system; hence, selecting the equipment array is a critical aspect for the system. A summary of scoring device possibilities is outlined in Figure 9 below.

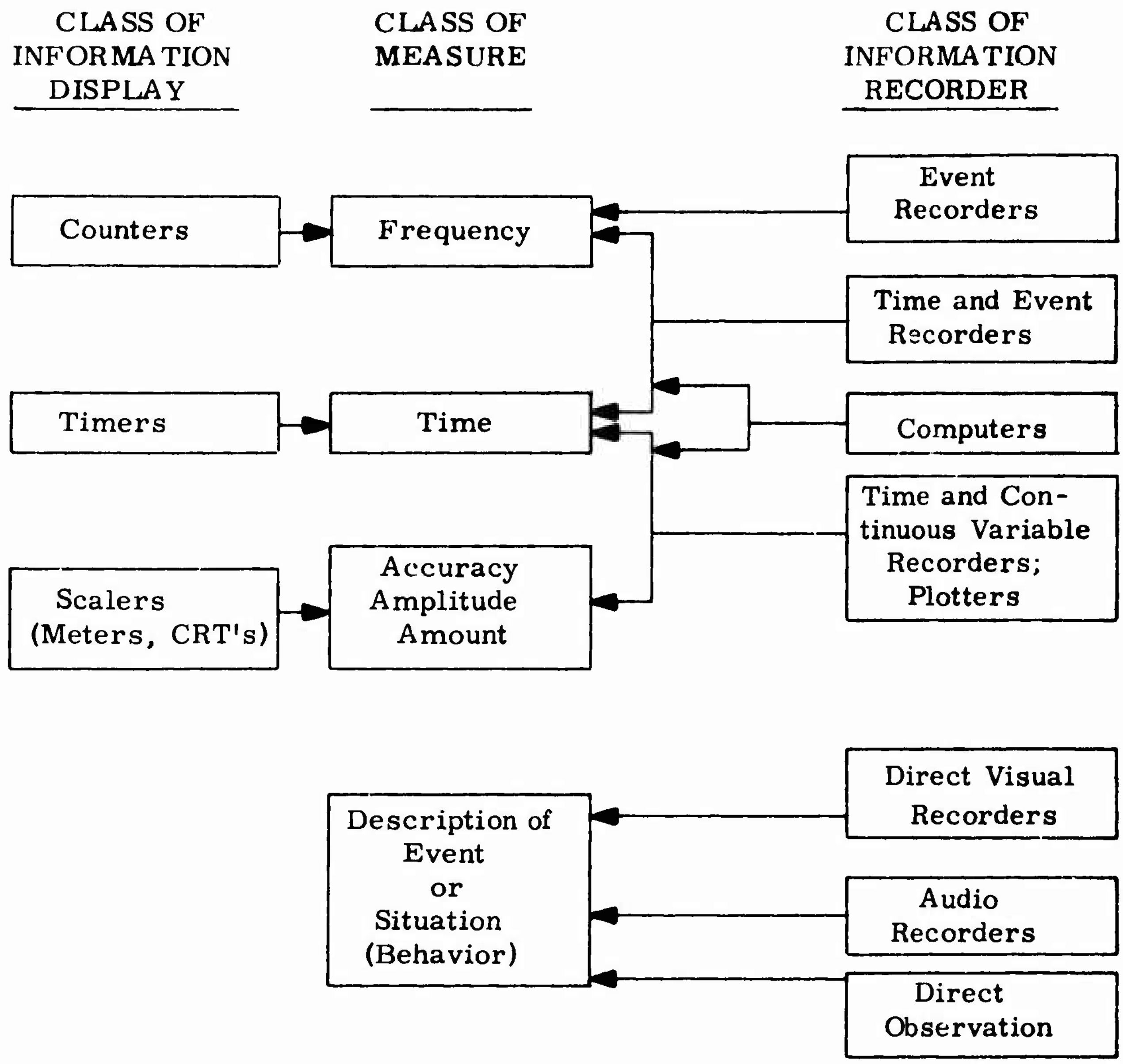


Figure 9. Data Collection Techniques and the Classes of Measures they Provide.

The computers used in aircraft control simulation will play an increasing role in performance evaluation because of their capacity to store output information from the trainee and compare this with normative data, and their capability to calculate rapidly, permitting immediate feedback of response or aspects of performance, duly weighted in terms of criticality per mission segment. To date, development of automatic

monitoring (scoring) systems for flight simulators has been hampered by several major obstacles. A recurring problem has been criterion determination, that is, the need for quantitative baseline values with which performance data may be compared for purposes of evaluation. Another major problem has been the determination of what measures best describe performance and what variables relate significantly to effective performance. These and other problems have hindered the achievement of automated performance measurement systems in simulators.

One of the first attempts to monitor pilot performance in the simulator was the UDOFT program (Sylvania, 1963). Using a general-purpose digital computer, this program demonstrated the feasibility of automatic performance monitoring.

An automatic task sequencing technique was developed in an Air Force study (Kurtzberg, 1963) which presaged a type of automatic programming for task scheduling in simulators. This type of task sequencing adapts to the present skill level of the trainee in determining the optimal training sequence as specified by training objectives (called adaptive programming when employing a digital computer). The purpose of the study was to investigate the feasibility of automating the instructor function of sequencing of tasks for presentation to trainees in flight simulators. Algorithms for task sequencing in real time were formulated for two classes of application: training of students for flight vehicle operation (operation teaching mode), and training for development of tactics skills (tactics teaching mode). Fourteen inflight emergencies were flow-diagrammed together with measures of performance and ranking of alternative responses available to the trainee. In the training sequence, tasks were automatically selected and presented, performance compared with established criteria and results recorded, and new task selection made on the basis of the trainee's previous performance. The logic of the technique suggests the possibility of automatic redirection of the training sequence as a function of the trainee's performance threshold of the moment.

Much current thinking is directed toward achieving a capability for automatic monitoring of pilot performance in synthetic ground training. An example of this trend is a program currently under development at the Aerospace Medical Research Laboratories at Wright-Patterson Air Force Base. The research effort is centered on developing a digital computer program for automatically monitoring human performance in the training simulator (Knoop, 1966). This research-oriented automatic monitoring program (called RAMP) provides an experimental tool for investigating the criterion problem as well as serving as an automatic performance monitoring system. In developing the automatic monitoring

program, it is assumed that no quantitative criteria exist. The problem is approached as a programming task, i.e., not accumulating data and processing it in a predetermined manner but rather establishing decision-making techniques to accommodate highly variable information on human performance. The computer program is designed to assist in the analysis and determination of performance measures and performance criteria, and using these criteria automatically monitors human performance. Part of the utility of the automatic monitoring program will be in the evaluation of criteria for flight tasks. These criteria need not be exact and may be easily altered. Of RAMP, Knoop says initially its effectiveness will depend on the accuracy and detail of the inputs provided by the user. "With the skillful application of dynamically programmed auxiliary routines, it can be made to resolve some of its own problems by iteratively collecting and processing performance data from many subjects."

In essence, the automatic monitoring program is designed to receive inputs regarding criteria, establish requisite matrices for consolidating these inputs, and then monitor the trainee's performance in order to test the criteria. The features of the monitor (implemented by dynamic programming) include: criteria analysis (already described); task sequencing (automatic redirection of the training sequence to provide individualized instruction); and automatic scoring. The intent of automatic scoring is to "free-up" the instructor so that he may perform observations of activities difficult to automate. The technique assumes the existence of a communication channel from RAMP to the trainee. Visual displays (digitally generated CRT) and auditory displays for this purpose are under consideration.

Another example, geared to operational requirements, is the research underway on an integrated monitoring program for evaluating the performance of advanced vehicle crews (Lincoln & Mangelsdorf, 1965). The objectives of this program are to develop an automatic system to assist in monitoring the performance capabilities and the physiological state of crew personnel, and to develop digital techniques for processing, displaying, and analyzing obtained data. Measures of critical body processes selected include the ECG, respiration, skin temperature, and blood pressure. Analog signals from the physiological sensors are sampled and the values digitized and recorded on magnetic tape (Control Data Corporation Model 160 computer is employed). The digital values are also displayed at a physiological monitoring console. Since the sampling rates produce tremendous quantities of data requiring excessive processing time and large storage for magnetic tape, data compression schemes have been developed to reduce data quantity to be transmitted while preserving the essential information in the signal. The performance measurement goal is the development of a battery of performance tasks to serve as independent predictors of aspects of flight operations. The battery of tasks

is a modification of that developed by the Lockheed Georgia Company (Alluisi, Hall, Hawkes, & Chiles, 1962) and is made up of tracking, drift monitoring, arithmetic, pattern comparison, and maze tasks. The presentation of these tasks is accomplished with the CDC 160 computer which generates digital signals that control the analog output of two digital-to-analog converters. The subject's responses are encoded and supplied to the computer, which prints out appropriate scores for each task. Computer generation of tasks permits task modification to meet changing requirements or add new tasks to the test battery without modification of equipment. The computer system also produces hard copy of test results, including a statement of the values established for the test parameters. The long-term goal of the crew-monitoring program is a system suitable for use in studies of human performance under defined stress conditions. Improvements envisaged include techniques for physiological measurement and increased capacity data processing equipment such as the SCAD system (simulation and control system) and the CDC 3200 computer (increasing the available memory storage).

By way of summary, however, little can be said on the scoring of pilot performance in the simulator except that the data are sparse and most of the studies are speculative. Similarly, research on measurement systems for the simulator consists principally of requirements and feasibility studies. It is our opinion, however, that the development of simulator scoring systems will receive considerable attention in the remainder of this decade, and the potential of assessing pilot performance in the synthetic ground environment will be fully exploited.

Research Issues:

A prime research requirement is the development of an objective and adequate evaluation system for scoring and assessing pilot performance. The literature indicates this to be a virtually untapped area for development. It also indicates that one can easily become dismayed by the extent and complexity of the problems in measuring and assessing pilot performance. One fact stands out: the effectiveness of training is seriously influenced by the effectiveness of measurement.

1. The basic question of how to measure pilot proficiency has yet to be satisfactorily answered. A number of theoretical writings and procedures for conducting measurement exist, yet in the operational environment the measurement picture is generally poor. The construction of accurate and valid performance measures for assessing pilot performance continues to be a recurring research requirement.

2. Establishing good criteria of pilot performance is a clear requirement. At present, the military is attempting to obtain performance standards by quality control in training. Army aviation, for example, is using the Pilot Performance Description Record to get measures from

which they can obtain estimates of change in a program. Field data on measures are needed as a means of eliminating irrelevant content from training, so as to answer more effectively the question, "Is the pilot being taught what he needs to know?" Without good criterion measures, there is no adequate way to determine how good the training is or how ready the pilot is.

3. An extension of inflight objectivity in scoring pilot performance is worth investigating. Several attempts were made during the 1940s to record objectively and permanently specific aspects of light plane flying performance by using graphic techniques (mechanical flight recorders depicting the effects of accelerations upon masses moving in certain planes) and motion photography (recording of flight instrument readings). These were used in conjunction with recording devices showing the movement of flight controls. The Friez Flight Analyzer, the Redhed Ride Recorder (Williams, McMillan, & Jenkins, 1946), and the R-S Ride Recorder (Viteles & Backstrom, 1943) provided continuous records of such features as airspeed, altitude, vertical acceleration, and elevator and aileron movements; and of such maneuvers as loops, Cuban eights, chandelles, slow rolls, and the falling leaf. Unfortunately, these techniques rated performance of aircraft as well as of pilot, hence confounded the measurement. Photographs were also made of aircraft instrument panels and actual manipulation of controls by the pilot (Viteles & Thompson, 1943; 1944; Wagner, Odbert, & Festinger, 1946), and special equipment involving photography was used in flight (Gardner et al., 1957; Jones, Milton, & Fitts, 1949). The graphic and photographic methods, however, are costly and time-consuming and require specially equipped aircraft. The records, of themselves, do not provide a measure of proficiency, but require techniques of evaluation which must wait until the processed film is available. Adequate reliability has not been established because of the small sample of raters used, and some questions have been raised concerning the relevance of the observed performance of flying proficiency.

The continuing development, however, of sophisticated, compact, lightweight recording equipment (e. g., videotape) will provide ways for more objective inflight scoring of aspects of pilot performance. For example, a program is underway at the Ohio State University for obtaining psychophysiological measures of pilot performance during flight (Billings, Eggspuehler, & Gerke, 1966). A method is being developed to assess a pilot's ability to cope with inflight stresses, i. e., correlating physiological variables with how the pilot actually performs. Coping behavior is measured in terms of demonstrable performance of the pilot under defined stress conditions. The investigators consider it essential that measures be selected which describe pilot actions resulting in some effect on the aircraft since these reflect an ability to cope with the environment (measures of psychophysiological status do not necessarily indicate this ability). In the initial work, a Hiller 12-E helicopter was instrumented

to monitor four system outputs: rotor rpm, positions of the throttle, collective pitch lever, and cyclic pitch control stick. The four outputs were recorded during critical portions of low-level electric powerline patrol flights on a 4-channel FM tape recorder. After each flight, the tape records were sampled and converted to digital tape format. The data were then condensed and plotted for visual inspection, and analyzed statistically, using an IBM 7094 computer. Indications are that this method of inflight measurement is capable of discriminating among pilots of varying experience and proficiency.

4. High payoff will accrue from a program of research for developing scoring systems for simulators. The eventual expectation is a fully developed, highly automated measurement system. Several types of studies are needed.

a. Feasibility studies to determine the range of practical problems, including hardware/device capabilities, cost factors, maintenance problems, and user acceptance.

b. Automatic scoring of the critical and important tasks in the pilot's job has been virtually unexplored. Thus, many cautions must be exercised in the eventual development of such a capability. Study should be devoted to determining representative events for observation, the applicability of measurement operations to task structure, and developing testing programs. In order to achieve the ability to predict how well the pilot performs a mission, baseline studies are needed. For example, knowledge of how the proficient pilots perform in the simulator can provide data for use in assessing trainee performance.

c. Since an automated scoring capability is desired which includes observing samples of behavior, recording, and processing of scores, and the readout of scores in usable form as a mission progresses, some effort should be devoted to specifying how best to instrument a measurement system. This should include the definition of hardware requirements considered in interaction with measures and behaviors, and tradeoffs between recording devices and observer involvement.

5. The Air Force standardization/evaluation check flights, although effectively employed, suffer from several weaknesses. These include the subjectivity of the technique, the one-ride concept which may yield atypical trainee performance, and the feature that the trainee is acutely aware of being tested. The reliability of the technique is not known. A small study is warranted, preferably within a unit of a selected command, to gather data on the reliability of the stan/eval check. Thought should also be given to developing ways for determining the relationship between cost/time and effectiveness of the method.

SIMULATION AND TRANSFER OF TRAINING

Simulation has assumed a formidable role in pilot training. Its use is based on the assumption that it provides a useful environment for the

learning of many aspects of the job. What is subsumed under this statement is a complex story, and much literature has been published about simulation and its various aspects. A number of summaries exist which extol the virtues or point out the shortcomings of simulators for training and other purposes, and discuss the philosophies of simulator design and usage for training (see Fraser, 1966; Gagne, 1962; Townsley, 1960).

Certainly, the greatest single gains in improving pilot training should come from the intelligent use of simulators in the instructional process. Logically, the case for simulation is quite clear. The increasing complexity of the pilot's job (including the requirements for advanced flight vehicle training) will demand the flexibility, control, and practice opportunities afforded by simulators. Also, imminent advances in design (e.g., automated data systems, adaptive techniques, increasing psychological fidelity) will serve to enhance the value of simulators for training. The rising cost of training in actual aircraft is another reason for espousing the cause of simulation, as is the increasing criticality of airspace utilization. Airspace problems are making mandatory a rigidly controlled profile-type training mission which allows little deviation for individual student progress during the mission. Unfortunately, the value of simulators for training is not well documented, nor have simulators been fully exploited. In essence, the data available today indicate that simulation is useful in the acquisition of flying skills but the usefulness is understood only in qualitative terms. Quantitative relationships are not precisely known.

The review of the literature in this section centers on studies of simulation employed in the training of pilots. Our objective is to assess the effectiveness of simulators in training pilots in terms of specific transfer effects to flying performance. Of prime interest are studies which compare the effects of simulator and nonsimulator training on subsequent flying performance. Researches which deal with the design of flight simulators are not of concern and are, for the most part, omitted. A few are included where the implications for pilot training are obvious. Studies which infer, rather than demonstrate, the value of synthetic devices are not pertinent to this review.

Two major topical areas are presented. The first reviews studies on the effectiveness of simulator training. The second reviews studies on simulation requirements for training and handles respectively, motion simulation, visual simulation, part-task trainers, and considerations on fidelity of simulation. Several literature areas pertinent to flight training were purposely omitted since they were more relevant to advanced flight vehicle training than to pilot training of the immediate future. These aspects emphasized long-term performance requirements in extra-terrestrial environments. Accordingly, studies on habitability (space cabin simulation), life support, long-duration work-rest cycles, confinement and isolation, and similar topics were excluded from the review.

Effectiveness of Simulator Training

Since World War II, flight simulators and other synthetic devices have been widely used for initial skill training of student pilots, transition training, refresher training, skill maintenance, and training for specific types of missions. During this period, considerable engineering and human factors design effort has gone into improving simulators to enhance their training capability. However, little information exists about their actual value for training. The research performed thus far indicates that simulators are useful adjuncts to flight training but has not systematically enunciated the conditions underlying the effective use of simulators for enhancing transfer of training. The studies contributing to this conclusion are summarized below.

Transfer of Training Studies: Evaluative studies of flight trainers, conducted up to the end of World War II (reviewed by Flexman, Townsend, & Ornstein, 1954), demonstrated that air training time could be saved by prior practice in a ground trainer, and that the proficiency of pilots trained in this way was frequently judged by instructors to be superior to that of pilots trained only in the air. From this base of information, a number of studies carrying into the 1950s attempted to extend and quantify information defining the training value of synthetic flight trainers for promoting flying skills.

A group of studies conducted by the University of Illinois sought to determine the effects of Link training in reducing the number of flying hours required to complete a private pilot curriculum. Williams and Flexman (1949) required students with no previous flight experience to learn to proficiency (three consecutive errorless trials) flight exercises composed of several related maneuvers. Students in the experimental group first learned the exercises in a Link Trainer and then relearned them in an Aeronca airplane. Control group students practiced only in the airplane. The experimental group learned the maneuvers to the criterion in the air with 28 percent fewer trials and 22 percent fewer errors than the control group. Thus, 25 percent of their flight training could be accomplished on the ground. Another study (Flexman, Matheny, & Brown, 1950) demonstrated that use of the Link Trainer, coupled with a set of specially developed training techniques (derived from a pre-experimental program designed to provide information on effective techniques for simulator utilization), could reduce the scheduled number of flying hours in a private pilot curriculum by more than 50 percent without loss of proficiency.

A number of transfer of training studies conducted with military pilots were principally concerned with evaluating the P-1 (SNJ) simulator for its effectiveness in teaching the flying tasks of the T-6 (SNJ) aircraft.

Mahler and Bennett (1950) found that the flying performance of simulator trained groups was significantly better than that of control groups trained only in the air, as reflected by less flight failures and accidents. But differences were not significant for check flight grades and for number of additional flights needed. Wilcoxon, Davy, and Webster (1954) found significant differences on objective measures of flying proficiency (radio-range and basic instrument procedures), also in favor of the trainer groups over the control groups.

Flexman, Townsend, and Ornstein (1954) compared the flying performance of a group given 40 hours of simulator training and 100 hours of T-6 training with a control group given the (then) traditional 130 hours of T-6 training with no simulator time. They found that the use of a simulator in Air Force Primary Pilot Training produced, in 30 hours less flying time, trainees who were as proficient on flight check rides as trainees given the normal number of flying hours. In addition, instructors' opinions indicated that the simulator-trained group was superior to the control group in overall flying proficiency. Attrition and accident data in both Primary and Basic indicated no adverse effects from the 30-hour reduction in T-6 flying time.

Performing more detailed analyses of the above data, Ornstein, Nichols, and Flexman (1954) found that P-1 simulator training was differentially effective for teaching the various maneuvers and components of T-6 flying performance. Comparisons of experimental and control group performance showed that the most effective simulator training occurred on those contact flight maneuvers heavily loaded with procedural components, and the least effective occurred on those items not well simulated by the trainer (advanced maneuvers). An unexpected finding was that P-1 training produced better instrument performance for the experimental group than for the control group who received instrument training in another type of instrument trainer. These authors conjectured that the effectiveness of the simulator as a training device could be improved by extending the range of simulation and by providing greater fidelity of simulation of certain critical components.

As part of a program aimed at determining the fidelity of simulation necessary to yield optimum transfer from a simulator to an aircraft, Matheny, Williams, Dougherty, and Hasler (1953), investigated the effects of varying control forces in the P-1 Link Trainer upon transfer of training to the T-6 aircraft. This study was specifically concerned with evaluating whether subsequent performance in learning climb and glide maneuvers in the T-6 aircraft was affected by differential amounts of control stick pressure used during previous training in a P-1 simulator.

Two groups of subjects were trained in the simulator, one with elevator control stick pressures roughly equivalent to those of the T-6 aircraft and the other with minimum control pressures. Both simulator groups learned the glide maneuver in significantly fewer trials than did the control group trained only in the aircraft. Differences between simulator groups were not significant. Performance of the simulator groups in the climb maneuver was also superior to that of the control group, but not significantly so. Thus, it appeared that the fidelity of simulator control stick pressures was unrelated to subsequent performance in the T-6 aircraft as determined by the trials to the criterion measure. The authors state that this finding corresponded to previous laboratory research which supported the hypothesis that in maneuvers of this type, transfer of training depends more upon a correspondence between the sequence or pattern of control forces required in the trainer and the aircraft respectively, rather than upon correspondence between the absolute amounts of control forces required.

Dougherty, Houston, and Nicklas (1957) evaluated the training effectiveness of four different ground training devices for teaching a large series of procedural and flight maneuvers. The devices used in ground training were: (1) an SNJ (P-1) operational flight trainer, (2) a procedures trainer, (3) a photographic mockup, and (4) the procedures trainer with an added tracking task. Four independent groups of subjects were assigned (one each) to each device for ground training. All groups then transferred to the SNJ (T-6) aircraft. Inflight performance was compared with that of a control group which received only inflight training. The subjects, who were private pilots transitioning to the SNJ aircraft, were given five learning trials on normal and emergency procedures in their respective devices. All trainer groups performed significantly better (fewer errors) on the first air trial than the control group who had no previous relevant learning experience. By the third air trial, no differences could be observed as a result of training with the different methods. The groups trained on the procedures trainer and the flight simulator showed the highest degree of transfer to the first air trial; but neither method was superior to the other. Both groups performed as well as the group which had already received five air trials. The conclusion was that normal and emergency procedures could be taught to transitioning pilots in a variety of ways and that for practical purposes, differences in performance disappeared after the first air trial. Another finding was that procedures could be learned as effectively on the ground as in the air.

Summary of Transfer of Training Studies: These studies represent, in large part, what is known about simulator effectiveness for training pilots. Before hazarding conclusions from these studies, it should be noted that generalizations from them are severely restricted because of variations in methods and practices. The studies differ in the measurement techniques used, the experience level of subjects (for example, flight trainees in one instance and experienced pilots in another), experimental and statistical procedures, simulator characteristics, point in time at which simulator training was given (for example, in some cases simulator training was given first, while in others a more or less alternated aircraft/simulator sequence was used), amount of simulator training given, amount of the task trained in the simulator (for example, in some cases, only two or three maneuvers were chosen as representative of all air tasks), and the background, experience, and qualifications of the instructors, and the way in which instructors were used (in some cases, simulator and flight instructors and evaluators were the same individuals).

In some studies (e.g., Flexman, Matheny, & Brown, 1950; Flexman, Townsend, & Ornstein, 1954; and Ornstein, Nichols, & Flexman, 1954) the use of a simulator was not the only innovation introduced into the training program. One study (reported in Flexman et al., 1954; and in Townsend & Flexman, 1954) used a specially developed training package to enhance the quality of pilot training. This package was especially designed "to be more efficient and effective than the conventional training program." The more important features of this package included: training the instructors in the principles of effective instruction; development of forms and methods for maintaining day-to-day control of the quality of training; use of knowledge of results, analysis of student errors and feedback; efficient use of training time; and specially developed training aids and devices. That this "special" course of instruction had significant effects on flying performance independently of simulator usage is documented by Townsend and Flexman (1954) who compared the performance of the control groups from two previous studies (Flexman, Townsend & Ornstein, 1954; Boyle & Hagin, 1954, already cited on page 56). In terms of "objective-type research flight checks administered by specially trained check pilots," at selected points during Primary Flight Training (at 18 hours, 60 hours, end of instrument phase, and end of primary), the special methods group was superior (mean error scores) to conventionally trained groups. These differences were significant at all measurement points except that at 60 hours. Since time and errors are usually highly correlated, one wonders whether the control group in the Flexman study (no simulator training) who made fewer errors than a conventional 130 hour group could also have attained some criterion performance level in less T-6 flying time than "normal" had they been given the opportunity.

Since the special methods training did provide a more efficient learning situation and better utilization of training time, it appears that the special methods training alone could have been profitably exploited to effect time reductions for T-6 training.

In spite of conflicting methodologies and other considerations of the studies reviewed in this section, there are certain consistent results which emerge from them, and some conclusions can be drawn:

1. Simulators and other synthetic devices have value for pilot training since they permit the learning of flight tasks on the ground which would otherwise have to be learned in the air. Thus, the training advantage is that they can reduce the number of actual flying hours required to achieve proficiency.

2. Transfer effects are greater for procedural tasks and other tasks requiring the patterning and integration of responses than for complex maneuvers. The indication is that procedures can be virtually 100 percent learned on the ground and can be taught in a variety of ways.

3. Less transfer to complex flight control tasks is obtained, presumably because of limitations in the degree to which simulator tasks actually represent the corresponding aircraft tasks.

4. There is no sound experimental evidence that simulator experience will produce more highly qualified pilots than training without it, or that simulators can be used beneficially in the absence of procedures for enhancing their use. Those studies where airtime reductions were achieved through simulator usage (e. g., Flexman, Townsend, & Ornstein, 1954; Williams & Flexman, 1949) employed the simulator as but one part of an overall program for enhancing pilot training, and procedures were developed for maximizing the simulator's contribution to training. In addition the simulator syllabus was structured to capitalize on its potential training contributions. Thus, obtaining maximum transfer from simulators may in many cases be as much a function of the way in which the simulator is used for training as it is a function of degree and fidelity of simulation.

Some suggestive data concerning transfer from a Link Trainer versus an aircraft used in the same way are provided by Ritchie and Hanes (1964). Subjects who learned instrument and contact flight tasks in "a relatively crude C-3 Link Trainer" before relearning them in the air required fewer air trials in both cases to achieve criterion than subjects without this experience. But in both cases, their performance was inferior to groups who learned the same tasks in a light airplane before transferring to another flight condition.

Simulator Usage for Training: How should simulators be used for training? Answers to this question are, at present, tentative since decisive quantitative data relating pilot proficiency to amount and sequence of simulator training are not available. Williams and Flexman (1949), in their first experiment with the School Link, sought to determine if preflight instruction in the trainer would reduce the number of hours of dual flight instruction normally required before solo in a light aircraft. Three groups of student pilots were given 4, 2, and 0 hours of preflight Link training respectively, with the expectation that the criterion of number of hours of dual instruction would be reduced as amount of simulator training increased. Flight instructors were permitted to solo a student when they felt he was ready. Separate instructors provided the Link and aircraft training. No differences as a function of amount of training were observed; however, instructors varied significantly in judging when the student was ready to solo. In the second experiment of this study, the same instructors were used for both Link and aircraft experience but the simulator usage was changed. Subjects now learned flight exercises to a criterion of proficiency (three consecutive errorless trials) in the trainer before learning them in the air. This study did show advantages for simulator training but probably because of the joint contribution of the simulator and the methods of instruction. Continuing the investigation of simulator usage, a study by Flexman, Matheny, and Brown (1950), mentioned earlier, was concerned with ways of using the trainer and with improving instructional techniques. The purpose of the study was to train students to private pilot proficiency in as few hours as possible by using any method deemed effective by the experimenter. A large number of instructional techniques and methods (e. g., intellectualization, knowledge of results, error analysis, etc.) were examined. All of these University of Illinois studies were concerned with developing methods for enhancing both the quality of pilot training and the value of the simulator for training.

Flexman, Townsend, and Ornstein (1954) applied this previously gained knowledge and experience to the conduct of an experimental pilot training program (see p. 143) for the Air Force. The contact portion of the T-6 Primary Pilot Training Syllabus was structured into 40 separate chronologically ordered flying proficiency lessons. Experimental group subjects then "completed" each lesson in the simulator prior to performing the same lesson in the aircraft. Thus, simulator training was alternated with aircraft training, and this procedure was shown to be effective for training students in the primary stage. Whether some other procedure might have been just as effective is not precisely known due to the lack of comparative data. However, owing to the experimental work leading up to the adoption of this particular usage of the simulator, it is assumed that other methods were tried and abandoned in favor of the alternated

sequence. Thus, it assumes a degree of validity for primary pilot training. That the methods of instruction (including the instructor) are also important in determining the value of the simulator should not be overlooked. Ornstein, Nichols, and Flexman (1954) have noted that a savings in airtime "is a function of both the simulator and the way it is utilized."

Thus, it appears that for training primary students an alternated sequence of simulator-aircraft practice is, at least, defensible. How much simulator practice should be given, however, is unknown. It appears that a "trials to criterion" measure for all tasks taught in present-day simulators would become excessively long and would not be justified since the amount of transfer to the airplane is definitely limited. Research should concentrate on the relationship between amount of simulator practice and "acceptable" or achievable levels of transfer to the aircraft. A key question concerns how much airtime is available for completing training. The answer defines the initial level of proficiency required for primary students and provides a starting point for in-the-air-training. Research aimed at maximizing transfer of training might be more profitably directed towards determining how much and what kinds of transfer can be obtained under known training conditions with a device of known characteristics.

Training experienced pilots to transition to new aircraft may not be favored by an alternated simulator-aircraft sequence. Dougherty, Houston, and Nicklas (1957) gave transitioning pilots five simulator trials on procedures prior to their learning to fly the T-6 aircraft. No deterioration of performance occurred on the first air transfer trial, and the group's performance equaled that of a control who had had five previous air trials. Thus, five simulator trials were equivalent in training value to five air trials for the procedures trained.

Fitzpatrick (1955) investigated the effects of different arrangements of aircraft and simulator training on the proficiency of crews transitioning to the C-124 aircraft. The design used permitted a comparison of different relative amounts of aircraft and flight simulator time and different sequences of aircraft and simulator practice. Three different "amount" conditions and three different "sequences" were used.

Amount Conditions

A1: 12-14 hours of aircraft training and 28 hours of simulator practice.

A2: 18-20 hours of aircraft training and 20 hours of simulator practice.

A3: 24-26 hours of aircraft training and 12 hours of simulator practice.

Sequence Conditions

S1: All or almost all simulator practice before aircraft training.

S2: Less simulator practice before aircraft training.

S3: Least simulator practice before aircraft training and "aircraft and simulator sessions were more or less alternated."

The proficiency of trainees was measured by means of "objective checks" administered in the aircraft and in the flight simulator. The findings of the study were:

1. Amount of simulator training time (and distribution of aircraft/simulator time) had no appreciable effect on measured pilot proficiency in the aircraft. Similarly, amount of aircraft time had no significant effects on simulator checks.

2. Sequence had only a slight effect on proficiency. The S2 sequence was apparently better for training emergencies and procedural tasks than the other sequences, although this superiority did not appear for total proficiency scores or for other subscores. No evidence indicated that alternated sessions were best.

3. Previous flying experience (both total flying hours and 4-engine flying hours) had no relationship to proficiency as measured by either the aircraft or simulator checks.

Thus, distribution of training time between aircraft and simulator for experienced pilots is apparently not an important consideration insofar as measured pilot proficiency is concerned. The sequence of instruction may or may not be.

Summary of Simulator Usage Data: It appears from the meager data, that for primary pilot training, the value of the simulator is enhanced when it is used in conjunction with a well-ordered training program and when simulator and aircraft periods are alternated. How much simulator training should be given is unknown. What to train is a matter for expert analysis and experimentation.

For training experienced pilots, an alternated sequence does not appear to be necessary, and it is also expected that less control over the training process is needed. While the limiting values are unknown, it appears that transitioning pilots can be given disproportionately larger amounts of simulator time (2 to 1) than aircraft time and still achieve the same level of proficiency as pilots given twice as much airtime (Fitzpatrick, 1955, referring to conditions A1 and A3). Thus, for the experienced pilot, the simulator may more readily substitute for airtime than it does for the novice, perhaps because the major portion of the transitioning pilot's task is the learning of procedures, and it has been demonstrated that procedures can be learned almost as well on the ground as in the air.

Research Issues: The simulator issue is perhaps the most frustrating aspect in pilot training. In some instances simulators are unduly praised, in others, they have been maligned to an extent not achieved by any other topic in training. There is no question but that simulation is a key factor in the training of pilots, and knowledgeable people look to it as a prime means for solving many of the training problems of the near future. What is perplexing is that the simulation lore concerning the value of simulators for training is extremely limited. The experimental evidence attesting to the value of simulators for training pilots is incomplete. Apparently transfer can be obtained under a wide variety of training and simulation conditions but it is virtually impossible to determine from the literature those conditions and combinations which best favor transfer. In essence, the research data demonstrate qualitatively the value of simulators for training. Beyond this, the data are inconclusive. This is unusual in a perplexing sort of way for this is an area where great sophistication (in training terms) should prevail because of the immense importance of the synthetic ground environment to pilot training. It is well understood that advances in simulation are sorely needed both technically and administratively. Yet, if one were to appraise the characteristics of the research environment, it would have these features:

Much of the research on transfer of training pertains to an earlier time, hence, is outdated in certain job conditions and problem perspectives.

Devices formerly employed as simulators bear little resemblance to the complex weapon system trainers and training facilities in use today. Still, little research has been published (or presumably accomplished) concerning use of these sophisticated equipments.

The research direction, emphasis, and sustaining power appear miniscule when considering the huge stake involved in the simulator

controversy. There is good reason to believe that the real issues are being neglected.

In our opinion, the research issue for simulation is simple but formidable: A heavy and sustained program is needed to gain an unequivocal understanding of the full value and meaning of simulation for aviation training. An intensive effort is also required for specification of the design requirements necessary for achieving the potential of simulation training.

Specific issues for research follow.

1. The manner in which simulators are currently being used for training in the Air Force should be determined by surveying the using commands. The survey should concentrate on the use of simulators for training pilots at all experience levels and for all types of missions. An attempt should also be made to determine how well such devices are doing their training job, taking into account the uses to which they are currently being put, and the uses for which they were designed.

2. Research effort is needed to clarify the relationship between different sequences of simulator training and transfer of training. It appears likely that an alternated simulator-aircraft sequence may favor initial pilot skill acquisition but that a block sequence may favor transition training. The effects of sequence should also be evaluated at early stages of skill acquisition. It is conceivable that the alternated sequence may provide the greatest gains at the very early stages of training but that a block sequence may be more efficient as experience with the task increases. Similarly, it can be expected that method and task will interact.

3. Relationships between amount of simulator training and subsequent skill acquisition should be evaluated. No meaningful experimental data exist for deciding how much simulator training should be given in order to achieve some specified criterion level in the air.

4. A special problem concerning evaluations of simulator effectiveness for training is in the statement of the criterion and the time at which transfer is evaluated. Several studies have attempted to demonstrate the value of simulator training by taking measures on some criterion of proficiency, such as mean errors in procedures. While in some cases of transfer of training it is conceivable that the first task effects will be greatest on the terminal levels of proficiency achieved on the transfer task, it is not necessarily the case that transfer effects will be manifested in this way. In most transfer of training studies it has been demonstrated that the effects of first-task learning are greatest at the beginning of the

transfer task where heightened proficiency on task components may be observed. However, at later stages of the transfer task, proficiency differences may no longer be apparent, and the net effect of the first task may have been simply to reduce the total amount of time required to master the transfer criterion of proficiency. That is, first-task training may more importantly affect rate of attainment of the transfer criterion than it does the absolute level of proficiency that can be achieved. Thus, some effort should be devoted to clarifying the relationships between these measures and the conditions under which they are relevant criteria for evaluating simulator effectiveness. For high fidelity simulators, proficiency measures may be preferred, but for lower fidelity simulators, training-time measures may be more informative.

5. Some attempt should be made to obtain evaluative data on the value of commercial airline simulators for training pilots and for proficiency maintenance. No studies were found during this review to indicate the value of these devices or the precise manner of their use.

6. Effort should be directed toward determining if simulators are being used for the purposes for which they were designed and how well they appear to be achieving these purposes. For example, devices designed to teach cockpit procedures should be used for this purpose and should be evaluated in terms of how well and under what conditions they best teach cockpit procedures as measured by transfer to the operational aircraft. It is suspected that there may be a degree of misunderstanding on the part of training personnel as to what certain devices can and cannot do and some attempt should be made to assess and correct this situation.

Simulation Requirements for Training

A number of researches have considered the question of what should be included in simulation for pilot training. Much of this research has been concerned with determining desirable design characteristics for maximizing transfer of training from synthetic equipment. In keeping with our purpose, we have minimized the review of research devoted to the discovery of design principles. The emphasis in this section is on the review of three groups of studies concerned with the requirements of simulation for training purposes: (1) investigation of transfer effects resulting from the simulation of relatively large classes of cues found in the flying environment (motion and visual cues), (2) evidence on the efficacy of part-task trainers, and (3) fidelity of simulation considerations.

Motion Simulation: The issue of motion simulation is a topic of much current interest. Although the research to date has not resolved the issues pertaining to the training value of motion in simulators, useful data are accumulating, the impetus coming partly from research on low-altitude, high-speed flight and from advanced flight vehicle research.

The literature reviewed here includes studies that (a) demonstrate that motion is a valuable and desirable feature of flight simulators, (b) suggest that the value of simulator motion is strictly a function of the transfer task characteristics, and (c) illustrate a very transitory transfer effect from motion training.

Early studies (such as Townsend, 1956; Muckler et al., 1959) indicated that motion simulation, particularly in early maneuver training was desirable. Similarly, pilot opinions, unsupported by experimental evidence, indicated that the value of a basic instrument flight trainer (the ME-1) was significantly enhanced by the provision of cockpit motion (Townsend, 1956).

A series of more recent experimental studies have indicated that simulated motion provides the trainee more cues than does the static condition and this enhances transfer performance. Besco (1961) found that motion cues facilitated precise tracking performance in the simulator (e. g., tracking in the pitch dimension in terrain contour flying). Buckhout et al. (1963) obtained evidence that simulated motion presented during the learning of a tracking task enhanced transfer of training. The Grumman Multipurpose Simulator which moves in three degrees of freedom: pitch $\pm 15^{\circ}$, roll $\pm 30^{\circ}$, vertical translation to $\pm$ 3 feet, and can produce 3-G accelerations, was used in all training and test trials. Three groups of twelve naive subjects were trained (12 trials) on a one-dimensional (vertical) compensatory tracking task under either a static condition, a 33 percent simulated vertical turbulence motion, or a 100 percent vertical turbulence motion. The criterion task consisted of closed-loop control in a vertically moving cockpit with 100 percent simulated vertical turbulence motion which the authors described as "flying a high-speed, low-altitude mission through clear air turbulence." (Flight was at mach 0.85 and under 500 feet). The subjects trained to track under the criterion (100 percent turbulent motion) made significantly lower error scores when transferred to the criterion task than did subjects trained under the static condition. However, there were no significant differences between the two motion groups. During criterion trials, the static group violated the vertical flight envelope (crashes or exceeding 500 feet altitude) 33 times, while the 100 percent turbulent motion group made only one violation.

Another study employing the Grumman Multipurpose Simulator (Ruocco, Vitale, & Benfari, 1965a; 1965b) investigated the effects of motion on performance. Subjects (Grumman employees with varying levels and recency of flying experience) were trained in a carrier approach and landing maneuver under static (no cockpit motion) and kinetic (cockpit motion) conditions. In the static condition, five groups, varying in conditions (n = 2 per group), were used; one group was used in the kinetic condition (n = 2). The performance of all groups was compared on criterion trials in the simulator with cockpit motion. Kinetic cueing during training significantly improved performance in the "transfer" task (i. e., the comparisons were between the static groups and the continuous kinetic practice group) in terms of percentage of successful landings, altitude error, time outside the flight path, and variability in smoothness of pilot control. The authors concluded that "kinetic cueing is a valuable and desirable adjunct to flight airborne simulation systems." Conclusions beyond this, however, were difficult to make. The study is complex and unwieldy, and is ambiguous in purpose (see pages 2 and 35 of the report) and in the treatment of the results.

While the above studies indicate that simulation of motion during training is desirable because it facilitates subsequent transfer performance, it must be noted that the criterion trials were given in a simulator and not in an aircraft. Since the validity of a simulator must ultimately be referenced to airborne transfer performance, the issue that arises logically concerns the relationship between simulator and aircraft performance. A number of studies have been conducted which have compared pilot performance in the air with performance both on fixed-and moving-base simulators. The results of these studies have shown that the importance of motion input in the simulator is directly a function of the type of task presented to the pilot.

Research performed by NASA (summarized by Rathert, Creer, & Douvillier, 1959; Rathert, Creer, & Sadoff, 1961) has compared test pilot's capabilities in flight over a wide range of steady-state and oscillatory conditions under all six degrees of freedom of motion with those on various fixed-and moving-base simulators. Comparisons made for piloting tasks included landing approach, longitudinal dynamics, longitudinal control, lateral dynamics, instrument presentation, and simulation of particular airplanes (variable stability research aircraft). The conclusions suggest that motion cues in the simulator are necessary only when they (1) contribute to improved control of the vehicle, i. e., help the pilot by supplying a necessary lead or anticipation cue, as in coping with a lightly damped or unstable vehicle, and (2) interfere with satisfactory performance, i. e., hinder the pilot in making a desired control motion as in using a very powerful or sensitive control system. Throughout the conduct of the studies leading up to these conclusions, a distinction was maintained between a

mandatory motion cue and a merely desirable motion cue. In the case of "landing approach," for example, it was noted that while motion inputs were not necessary, this did not mean "that the pilot would not like motion or use it if given to him. It means that he can get by without it." (Rathert, Creer, & Douvillier, 1959).

On the basis of the researches performed by NASA it would seem that motion cues in the simulator are necessary only in specific cases with the decision to include motion being based on the control dynamics of the particular aircraft being simulated. These conclusions, however, are based on the performance of highly experienced test pilots whose data are not representative of the general population of pilots. Additional evidence concerning the role of motion for enhancing transfer of training to operational flying tasks is provided by studies of helicopter training.

In a simulator study of hover training in a helicopter, Feddersen (1961) trained an experimental group of subjects in hovering on a six-degrees-of-freedom simulator and trained a control group under static conditions. Upon reaching an asymptote in training, each subject was given six, 2-minute hovering trials in a helicopter. As expected, the motion cue training group performed better initially in the air than did the static training group but the differences disappeared by the end of the six-trial flying session. The case for motion simulation is lessened somewhat since the original difference in transfer effects between the two groups was not maintained at the end of the criterion task session. The important feature is that the greater initial transfer of training may be explained simply in terms of familiarity with a larger number of task aspects. Because the difference disappeared so rapidly, Feddersen considered that the use of a motion simulator was difficult to justify.

A recent study (Caro & Isley, 1966) conducted for the Army involved an evaluation of the effects of high-fidelity motion simulation on subsequent transfer of training. Helicopter pilot trainees were given 3 1/4 or 7 1/4 hours of presolo training in a synthetic contact helicopter training device. Their subsequent performance was compared to that of both a conventional and a "blind" control group. The unique device, the Whirlymite Helicopter Trainer, is a one-man helicopter mounted on a ground effects machine through an articulated linkage which allows freedom of movement in six dimensions and preserves the handling characteristics, and visual, auditory, and proprioceptive cues of the inflight task.[21] No differences were observed between Whirlymite-trained groups as to the amount of training, but training with the device resulted in 20 percent fewer attrition cases than

---

[21] Described by Caro, P. W., Jr. Reduction of helicopter pilot attrition through synthetic flight training. Paper read at 73d Annual Convention of the American Psychological Association, 3-7 September, 1965.

for the control groups. Insofar as flying proficiency was concerned, device-trained groups acquired the skills necessary to operate the helicopter safely in solo flight with significantly less inflight training than did the controls. They also performed in a more satisfactory manner as reflected in presolo daily grades while acquiring flying skills. There were no significant differences between experimental and control groups on any of the flight performance measures after the first solo flight. Despite exposure to the full range of helicopter motion, transfer effects did not last beyond solo. This study represents an important methodological contribution to motion research since the "ground" training device used represented task conditions that were highly realistic to the flight environment, i.e., a tethered small helicopter. No other studies were able to achieve such high phenomenal equivalence between training and test.

The preceding two studies strongly suggest that the value of motion simulation for training is limited. While initial benefits do accrue from relatively small amounts of motion training, they dissipate very rapidly as the control groups acquire experience with in-the-air motion. High-fidelity motion simulation, however, may have significant value for reducing attrition in subsequent flight training. It is interesting to compare these results with the results of studies which evaluated the light plane as a selection and training device (pp. 56-58). Virtually identical results were obtained, i.e., small initial training transfer gains but significant results as to elimination of motivationally weak students.

On a logical basis, one might expect that the importance of motion cues in flying might increase as the experience level of the pilot increases and as he "learns" the meaning and implications of certain types of motions in the air. Flexman,[22] reports that in a study performed by the Air Force Personnel and Training Research Center, students and instructor pilots had differing experiences when they attempted to make an instrument takeoff in a T-6 instrument trainer that did not use a motion platform. Instructors tended, on their first flights in the trainer, to spin-in shortly after takeoff, but none of the student pilots had this problem. Conclusions were that the students were mechanically cross-checking their instruments, as instructed, but the instructors were not. Presumably, the instructors were attempting to rely on associations between motion cues and responses learned in the actual flight situation which cued them to look at certain instruments when they felt certain motions. Apparently, the instructors got into trouble because of the lack of the initiating motion cues in the trainer.

---

[22] Flexman, R. Man in motion. The Connecting Link, 1966, 3 (1), 12-18. (General Precision, Inc., Link Group, Binghamton, N.Y.)

The data suggest that motion cues (proprioception) become more important as experience with the flying task increases. Apparently, pilots learn to rely on these cues and use them as preconditions to action. However, alternative sources of information were available in the above case and did adequately substitute for lack of motion for the students. Again, the distinction must be raised between mandatory and desirable motion simulation. Should the pilot be trained to rely on motion cues? Research addressed to this question must also consider cases in which high time-sharing requirements (as, for example, in low-level, high-speed flight) are also imposed on the pilot. Motion cues may then assume more prominence in directing his responses since, in effect, alternative information sources are not readily accessible to him.

Research Issues: The value of motion simulation in training is not yet resolved. There are many difficult problem areas, beginning with the basic question of whether motion should be provided at all. Some researches have suggested that in most cases motion cues are not strictly necessary, but other studies have demonstrated that they are, at least desirable. Whether motion cues are merely desirable, or should be mandatory, is difficult to determine. An essential question that must be asked, however, is: Does more effective transfer of training occur when motion is provided than when it is not? If the answer is affirmative, then increased training gains should be weighed against other criteria (cost, safety, maintenance, etc.) as a basis for deciding whether to include motion in the simulator. This evaluation should be made somewhat independently of the question of whether or not motion is (strictly) necessary. Pushed to the extreme, of course, one must conclude that simulators per se are not mandatory for training, but there are valid reasons concerned with training efficiency for using them. The value of motion simulation in training is still largely unknown and some effort should be directed toward clarifying this area. Similarly, if motion is to be provided in the simulator, then effort is also needed to clarify what kinds and what amount.

There is some evidence to indicate that the importance of motion cues may be a function of the experience level of the pilot. But whether motion in the simulator is more important for initial skill acquisition or for training experienced pilots is not known. Briggs and Wiener (1959) and Flexman (see p. 156) suggest that motion cues are, at least, more used and more relied on as experience increases. Muckler et al. (1959), however, suggest that motion is more important at early stages of training, particularly insofar as it interacts with contact cues. Research

may be able to resolve the question of when motion training should be given, but the problem appears to be still more complex and involves distinctions between learning and performance and the functions of motion cues.

That pilots do use motion cues (at some level of awareness) in controlling their aircraft is probably a valid statement. The function of these cues is, however, unknown. Ruocco et al. (1965) suggest that motion functions as a "general alerter" and, hence, serves to alert the pilot to a changing state that prompts increased attention to visual cues. Elsewhere they qualify this statement (pp. 84-85) to include the possibility that in at least certain cases, motion may provide specifically useful information that the pilot uses directly for controlling his vehicle. Feddersen (1962) suggests that motion cues serve to quicken the pilot's entire response network. More research is needed to clarify the function(s) of motion cues.

Experienced pilots apparently use motion cues for some purpose; that is, motion is somehow sensed and used in the performance of flying tasks. In some respects the fact that motion cues are used by the experienced pilot is not clear justification for including motion in the simulator. The test of value of a simulator for training is whether practice in the device results in transfer of training to the operational tasks. Thus, the question for research is: Does motion simulation affect the learning and transfer performance of pilots trained with it? The desirability of attempting to train pilots to rely on and use motion cues is a separate question. Whether motion cues become a more important source of information as the pilot's workload increases is also an open question.

Extra-Cockpit Visual Simulation: Simulation of the external visual environment is primarily an engineering design issue, and much research has been devoted to determining what contact cues to simulate. It is obvious that complete visual simulation is not necessary, nor is visual simulation required for every aspect or segment of the mission profile. The essential problem is that the factors to be simulated and their weightings are not precisely known. Also, what to simulate is a relative issue in the mission context, for the appropriate external visual environment changes as a function of the task/segment in the mission profile.

The studies reviewed here are limited to the research on the training of pilots in the use of visual cues, i.e., not on training in the use of vision but on how the visual (contact) world is represented in

flight. Two types of studies are included. The first is concerned with demonstrating the training value of external visual (contact) displays for training pilots and the conditions which govern the effective use of these displays for training. The second is concerned with training the pilot to divide his attention among the sources of information, both extra- and intra-cockpit, that must be attended to and the activities that must be accomplished while flying.

Man's visual capacities are largely "givens" in our context and are not considered as direct training problems. Visual capacity is of primary interest to basic research and, somewhat, to human factors design. Representative problems for visual research are the determination of what cues are used in contact flight and how they function, both within given mission segments and in the overall flying regime. Thus, problems such as relative motion and dynamic visual acuity are of importance. The problems for human factors design include the qualitative and quantitative translation of relevant visual cues into physical entities to be displayed in a manner approaching the perceptual equivalence of real-world cues. Together, the solutions to the above problems define design requirements for developing displays which provide the trainee information relevant to the learning of contact flight skills in the simulator. For information on the data and problems associated with the requirements and design of visual displays for simulating contact cues for pilot and astronaut operations, the reader is referred to studies such as: Buddenhagen, Johnson, Stephan, and Wolpin (1963); Buddenhagen and Wolpin (1961); Gerhardt and Johnson (1963); Gibson (1950; 1955); Pfeiffer, Clark, and Danaher (1963); and Whittenberg and Wise (1963). Given that a contact display has been developed, the value of the experience from using it is the point of departure in the present discussion.

Training on Contact Cues--The value of a contact display for training student pilots to land an aircraft was demonstrated in an early study performed at the University of Illinois. A relatively crude contact landing training device, a manually rotating blackboard for use in teaching ground reference maneuvers, was shown to have substantial training benefits for training private pilots when used in conjunction with the School Link (Flexman, Matheny, & Brown, 1950). The device, however, was not effective for training military pilots (Ornstein, Nichols, & Flexman, 1954) when used with the P-1 flight simulator. Apparently the "goodness" of this device was also, in part, dependent upon the type of instruction given with it.

A study by Payne et al. (1954) used a more elaborate landing display in conjunction with the SNJ Operational Flight Trainer. The display

used was a closed-loop projection system. The runway image on the screen in front of the pilot changed with changes of the simulated aircraft with respect to runway position. Private pilots trained with the simulator and the contact device took 61 percent fewer air trials to reach proficiency and made 74 percent fewer errors in landing approaches than control group subjects.

Creelman (1955a) used the same contact landing display and a "Cyclorama" visual display in conjunction with the Link SNJ Trainer in an operational flight training program for the Navy. One group of subjects (n = 15) received flight training in the trainer, with contact landing simulation, prior to flying the aircraft. A second experimental group was shown films of contact landings and the runway image simulated by the landing device. This second group did not fly simulated landings with the trainer, however, and therefore did not make the motor responses associated with specific runway configurations. This group served as a control against the possibility that the value of the contact landing display was primarily perceptual and had little to do with actually making the flight control responses in a closed-loop relationship with the display. A control group with no pretraining received air practice only. Criterion performance in the aircraft was defined by instructor ratings. Creelman concluded that the performance of the group which actually flew the flight trainer with the contact landing display was distinctly superior on landing performance in the overall training program. Differences could not be accounted for by the intellectual training given the other experimental group. Apparently, the trainer's psychomotor aspect was the key to the effectiveness of the procedure in contrast to stimulus training alone.

Adams and Hufford (1961) sought to determine if perceptual pretraining on contact landing cues would influence the subsequent acquisition of contact landing skills in a trainer. Naive subjects were shown visual scenes of runway landing patterns (programmed sequences of the trainer's visual display) as they would appear if the subject were actually flying an SNJ aircraft in a night landing exercise. During the pretraining session, the subjects were required to make judgments about correctness and incorrectness of the landing patterns presented,[23] but were not required to fly the simulated aircraft. Transfer criterion performance was measured when the subject actually flew simulated contact landings with the flight trainer. Performance was compared with control subjects who

---

[23] Cf. with studies of "stimulus predifferentiation;" e.g., Vanderplas, 1958.

had not been given perceptual pertraining but who simply learned to fly the landings in a closed-loop relationship with the display. All criterion trials were in the Contact Analog Landing Research Tool (Device 20-L-10a), which simulates an SNJ aircraft and the contact runway cues of a night landing pattern. The perceptual-verbal pretraining had no apparent effect on the subject's subsequent learning of the contact landings in the trainer, where he was required to make control responses in conjunction with the presented visual scenes. Where differences were observed, they generally were in the direction of inferiority for the experimental group. Several conclusions are possible. The most obvious one is, of course, that open-loop training is simply not effective for teaching an essentially real-world, closed-loop task. Apparently, a contact display by itself has little training value. The subject must make control responses to the contact cues in a closed-loop fashion. Other conclusions that may be advanced are that insufficient training was given or that the method of training was inappropriate for the task. Requiring the subject to make some kind of directional motor response to the visual presentation rather than the simple verbalization of correctness or incorrectness might have been a more effective training method. Adams and Hufford interpreted their finding, however, to support the existence of a learned interaction factor (see p. 165) between the perceptual and motor components, which could not be learned with the programmed presentation. The pretraining exercises might have given the subject a tendency to be unduly preoccupied with the visual scene and, consequently, he failed to timeshare his scanning of the contact world adequately with critical cues on the instrument panel.

Timesharing Training--Several studies have investigated the use of contact displays for improving timesharing behavior of pilots. The timesharing issue begs the question of what contact cues are attended to by the pilot in performing his flying tasks, and also whether external visual displays, and of what kind, should simulate this or that aspect of the flying task. It is, rather, concerned with teaching the pilot to divide his attention among flying tasks and to develop effective scan patterns for sampling cues from both intra- and extracockpit visual sources. Also germane is the question of whether flight control and visual tasks should be practiced concurrently in training.

Pfeiffer, Clark, and Danaher (1963) gave training in timesharing between intra- and extracockpit visual cues and between the visual cues and flight control tasks. Ten experienced military pilots served as subjects; half were jet qualified and half were qualified in multiengine propeller aircraft. Pilots were trained in an F-100 fixed gunnery trainer (cockpit and controls of the F-100A aircraft) with an external visual

display which consisted of a large hemisphere onto which a horizon could be projected as well as target aircraft. Three specific aspects of timesharing behavior were examined:

1. Inside cockpit scanning (cockpit emergency detection).
2. Inside-outside cockpit scanning (formation flying).
3. Outside cockpit scanning (intruder detection).

Subjects were specifically informed that one of the purposes of the study was to investigate the training potential of the simulator with respect to timesharing. The overall results of the study indicated improvement in simulator performance with practice, with respect to both emergency/intruder detection (as measured by time to detect) and aircraft control (as measured by altitide holding ability), for pilots with varied flying experience. The authors concluded that timesharing behavior did improve with practice in the simulator; that overall aircraft control behavior improved; that training in timesharing could be installed; and that the flight simulator with a nonprogrammed (i.e., closed-loop) visual display is a "promising technique for implementing such training."

A recent study (Gabriel, Burrows, & Abbott, 1965) was conducted to provide a more extensive test of the effectiveness of training in timesharing as well as to determine if a simplified device could be used for such training. Sixty Marine pilots divided into upper and lower experience groups participated in this experiment. The experimental group received eight hours of timesharing training in a simulator plus approximately four hours of tachistoscopic instrument speed reading training. The training device which the pilots "flew" was a simple, generalized cockpit which included basic flight instruments and also signal lights for signaling intracockpit malfunctions and extracockpit events (intruders). Experimental group performance was compared with that of a control group who had received no prior timesharing or tachistoscopic training. Comparisons were made on a series of criterion flying tasks in the A-4 Operational Flight Trainer (Device 2F76) with a visual attachment (programmed film sequence). The timesharing training group showed significantly greater ability to detect outside-the-cockpit emergencies (i.e., intruders) than the control group. This improvement was achieved without compromising other flight tasks. Timesharing training was effective at both low and high experience levels. The tachistoscopic training was effective both in increasing reading speed and accuracy and in decreasing the time required to scan the instrument panel.

Thus, the preceding studies suggest that timesharing training is needed by pilots at all experience levels (since such training improves

the detection of extracockpit emergencies). Gabriel, Burrows, and Abbott demonstrated, however, that it can be given with less costly and less complex devices than full mission simulators employing external visual attachments. Whether it is necessary to include pilot control tasks in response to visual cues as part of the training routine is not clear. The answer to the basic question of whether contact cues should be simulated appears to be yes; but more definitive work is needed.

Research Issues: Since external visual (contact) displays are only a portion of a total training system, the effectiveness of these devices may be as much a function of the fidelity of the simulator with which they are used as it is a function of their own fidelity. Research is needed to clarify the training effectiveness of contact displays, both independently of, and as a function of, the rest of the simulator system, with a view to determining how simulation requirements (and training value) change.

Motion cues in the simulator, which may or may not affect the learning of procedural and flight control tasks, must certainly affect the training value of contact displays, and effort is needed to clarify these relationships and to find satisfactory ways of presenting contact and motion cues in relation to each other for effective training. It is not known, for example, if moving the display, or changing the relationship of elements within the display, or moving the aircraft in response to pilot control inputs are phenomenologically equivalent to each other. A related question concerns the value of open-loop training for closed-loop contact tasks and the conditions under which it may be effectively used.

Effort is also needed to clarify the relationship of pilot experience to the required fidelity of simulation of contact displays. Apparently, for student pilots at least, even very crude contact displays have considerable training value, provided that "good" instruction is given in relation to the device. Quality of instruction may substitute for absent contact cues in many cases.

The use of simulators with external visual attachments for time-sharing training appears to be an inefficient use of complex and costly equipment. Simpler devices should be explored as to their effectiveness. Determination should also be made of the value of tachistoscopic training for improving instrument reading skills so as to allow the pilot greater time to attend to extracockpit cues.

Part-Task Trainers: This discussion of part-task trainers centers on the value of these devices for pilot training, specifically, the effects of part-training on subsequent transfer to the larger complex of flying

tasks. The value of part-task trainers has frequently been discussed within the more general framework of the classical part-whole controversy on learning efficiency as a function of practice schedules. Although these conditions for learning have been studied for many years, few unvarying generalizations have arisen because of the complexity of the interacting variables: the nature of the learning, the characteristics of the learner, and the conditions of practice. The majority of the part-whole studies have been carried out in the laboratory with such simplicity of task structure (verbal learning, simple motor tasks) that a correlation with the scheduling of practice requirements of flying tasks cannot be made. The relevancy of classical part-whole research data to training in part-task devices is seriously restricted because these devices do not provide practice on components of a total task in the classical sense. Rather, they provide practice on whole units or sequences of skilled behavior which are learned and transferred intact to the total job of flying an aircraft, where the learned skills must now be performed in conjunction with other, different tasks. Consequently, little can be gained from this literature that is cogent for our purposes. Detailed accounts of the part-whole issue and compilations of the shortcomings in the research data have already been prepared, (e.g., McGeoch and Irion, 1952; Naylor, 1962; Osgood, 1953; and Van Cott, 1955).

While part-task trainers have utility for pilot training (e.g., Dougherty, Houston, & Nicklas, 1957; Miller, 1960; Parker & Downs, 1961; and Pomarolli, 1965), one criticism that has been leveled at them is that the trainee does not learn properly to timeshare the skills learned in the part-task trainer with other skills that he must use interactively in the mission environment (Adams, 1957; 1960; and Adams & McAbee, 1961). The implication for flight training is that the skills learned by the pilot in a part-task trainer require additional practice in the complete task context before proficiency can be maximized on the total task because of the necessity for learning timesharing between the tasks. Two studies conducted at the University of Illinois investigated the timesharing hypothesis.

Adams, Hufford, and Dunlop (1960) trained two matched groups of private pilots in a hypothetical "toss-bomb" maneuver. One group received separate training on the flight control and procedural parts of the maneuver; the other learned both classes of tasks concurrently. After equal amounts of training, both groups performed the maneuver in the "whole-task" version (the criterion task was simply continued practice for the second group). All training and criterion trials were given in the SNJ OFT and were "under-the-hood" instrument flights. Analysis of the criterion task showed that for the procedural activities,

the performance of the part-task group was significantly inferior on the first trial to that of the whole-task group. The flight control tasks were performed equally well by both groups on the first whole-task trial. The authors accepted the findings as supporting the hypothesis of a learned interaction between flight control and procedural tasks. The part-task group did not have the opportunity to acquire the interaction and timesharing in training. Consequently, when the requirement for timesharing performance of the two classes of responses was imposed, the group trained on "parts" was inferior to the group trained via the "whole" method. It is important to note, however, that the "inferiority" did not carry over beyond the first "whole-task" trial in the simulator.

A second study (Hufford & Adams, 1961) investigated the contribution of part-task training to the relearning of a whole task (see also p. 174). Subjects from the previous study (above) were recalled after a ten-month interval of no practice and required to relearn the same maneuver. Half of the subjects were given part-task refresher training (ten trials) on procedures. The others were given only whole-task training. The part-task group regained proficiency by the third whole-task trial, and the control group (whole-task training) relearned the maneuver in five trials. The value of part-task training for proficiency maintenance was demonstrated since it reduced the number of trials required to regain whole-task proficiency. Since the part-task group did not exhibit complete proficiency on the first criterion task (whole-task) trial, the results were taken as support of the timesharing hypothesis. It was concluded that the use of part-task trainers should be followed by some integrative whole-task training to allow the trainee to regain the apparently lost timesharing skill.

It is of interest to note that the original learning acquisition curves (Adams, Hufford, & Dunlop, 1960) for the two groups show a consistent superiority in learning for the part-task method. The finding of no significant differences after the first criterion trial lends weight to the notion that there is considerable training value to be realized from properly used, relevant synthetic devices, whether they be part-task or full-mission simulators. The findings in studies of this type that transfer is less than perfect should not necessarily be viewed as a limitation of the device, but rather as a limitation on the way in which the device should be used within a training program so as to enhance its transfer value. Part-task trainers, it seems, are most useful for both initial skill acquisition and for refresher training, when practice in the part-task trainer is followed by a period of practice during which the trainee is given opportunity to integrate the learned skills into the pattern required by the transfer task.

Research Issues: There is a paucity of data comparing transfer effects from part-task trainers to those from more complex simulators, and some effort should be devoted to determining the relative contribution of each to training. This should be tied in with the growing interest in backing off from high-fidelity simulation to reach acceptable levels of fidelity for achieving transfer of training, with accompanying cost reductions (see p. 167). There is a suggestion in recent literature (Pomarolli, 1965) that certain types of part-task trainers (instrument trainers) may be doing as effective a training job as more complex simulators when small anticipated gains are weighed against increased costs. Data are needed on this point.

The major consideration of this section, that part-task training does not lead to optimum first-trial transfer levels when a whole task is performed by the pilot, should be further explored. Future research in this area might profitably investigate how quickly and effectively the integration of different skills learned in part-task trainers proceeds. One outcome would be a family of curves depicting the number of trials required to integrate different numbers and kinds of skills. Another outcome concerns the issue of whether total training time can be reduced by the judicious use of part-task trainers plus integrative training over that required by whole-task training alone.

Fidelity of Simulation: Considerable literature exists on the topic of fidelity of simulation. For the most part, these studies have been addressed to improving the design of synthetic training equipment in order to maximize transfer of training to operational tasks. Strictly speaking, they deal with design problems rather than training problems, and hence, do not fall within the purview of this report. However, design and transfer of training are intermingled in the sense that fidelity is evaluated in terms of the performance of subjects trained on devices of a given design and fidelity of task representation. For this reason, it is worthwhile to cite briefly the problem of fidelity of simulation for training. Previous reviews of fidelity of simulation, e.g., Muckler and his colleagues (1959), have noted that definitive answers to fidelity problems have not been provided by the experimental literature. Factors cited include the lack of directly relevant transfer of training studies, inability to generalize from the studies that are available, inability to define the pilot's job precisely, inadequate measurement techniques, and lack of generalizability because of oversimplified laboratory tasks. Muckler has also delineated a large number of research needs within the area of fidelity of simulation. For the most part, these needs are still current.

Current practice in simulator design attempts to achieve high engineering fidelity between aircraft and simulator task elements. This approach is "forced" on pilot training because of the many problems in fidelity of simulation for human use. Thus the traditional belief is that the closer the training device resembles operational equipment, the greater will be the transfer of training. This reasoning results in the development of extremely costly training equipment. Much has been written on the issues and problems of fidelity of simulation. There is considerable evidence, however, that deliberate deviations from fidelity of simulation may lead to higher levels of transfer than does exact simulation. Thus, the weaknesses in current simulator design and mounting costs dictate the need to modify current practice.

A research program (Demaree, Norman, & Matheny, 1965) is currently being conducted for the U.S. Naval Training Device Center using the UDOFTT (Universal Digital Operational Flight Training Tool; see also AMRL-TDR 63-133) to investigate required degrees of fidelity of simulation for transfer of training. Plans are to program degradations into the system and to assess the effects of the fidelity reduction on performance under full simulation. The authors state, "With the increasing complexity of weapon systems, the rising cost of providing high simulation fidelity makes it imperative that such studies be conducted in an effort to determine acceptable levels of reduction in fidelity with accompanying reductions of OFT cost." Thus, cost considerations are forcing a reevaluation of the practice of making simulators as much like the airplane as possible and can be expected to lead to the development of simpler devices whose training value must be determined. Because of cost factors, one may also predict a resurgence of interest in fidelity of simulation problems and research to solve them.

Research Issues: The reader interested in research issues for determining "optimum" design characteristics of simulators for "enhancing" transfer of training is referred to Muckler, et al. (1959). Many problems exist, and it can be expected that answers to these problems will become more urgent if cost considerations force simulator design to lower levels of fidelity than currently provided.

There is some suggestion in the literature that fidelity requirements may be different for individuals of different skill levels. Thus, the required fidelity of simulation in a device must be conditioned by the tasks that are to be taught through use of the device, and by the levels of experience (or skill) of the trainees. How fidelity considerations interact with other instructional variables should also be examined. The quality and style of instruction required, for example, to meet

stated training objectives, as well as instructor workload, are apparently affected by device characteristics. Apparently, more and better quality instruction is needed when fidelity is low (Cox, Wood, Boren, & Thorne, 1965; Dougherty, Houston, & Nicklas, 1957). These relationships should be ascertained.

MAINTENANCE OF FLYING PROFICIENCY

Loss of flying skills as a result of disuse is of considerable concern in Air Force operations. There are instances where availability of resources preclude all pilots getting enough opportunities (flying sorties) to sufficiently exercise their skills. Thus, staleness through disuse, or aircrew not being "peaked" in proficiency brings up problems of skill maintenance and retention. The questions of importance for flight operations are: What are the effects of various no-practice periods on flying proficiency? What is the threshold of the dangerous no-practice time interval? What aspects of flying skill degrade most significantly through disuse? What procedures are available to keep pilot skills peaked?

Retention of Flying Skills

Answers to the above questions are differentially available in the literature. While the body of research dealing with retention, forgetting, memory, skill maintenance, and the like, is substantial, it is centered for the most part on laboratory situations employing simple task situations. These are poorly aligned in pertinence to the above questions and very little of the data can be generalized to the job of flying. Much of the research has been concerned with verbal learning (poetry, serial lists, nonsense syllables, etc). The perceptual-motor tasks have essentially been simple (mirror tracing, ball tossing, psychomotor tests, mazes, etc.) but have included more complex performances such as playing the piano, typewriting, Morse code transmitting and receiving. The scholarly presentation of McGeoch and Irion (1952) is recommended for an appraisal of the history and problems of retention and forgetting research up to about 1951. In recent years more complex motor tasks have been employed in research (e.g., complex tracking tasks). The 1961 review of long-term retention of learned skills by Naylor and Briggs (1961) is additionally recommended for an assessment of motor skills research.

The studies in the literature are essentially of two classes. The first class is concerned with the availability or continuance of skills as a function of disuse over time. Here the interest is in retention as a function

of the type of task used (verbal, motor, discrete, continuous) and the length of the interim time period. The second class is concerned with the variables that affect retention and includes such factors as amount and degree of original learning, type of learning technique, nature and amount of interpolated activity following original learning, the duration of the interim period between learning and test of retention, procedural manipulation, organismic manipulation, and the method by which retention is measured.

The state of information, so far as pilot training is concerned, suffers on several major counts. Much of it is based on simple verbal learning studies which are irrelevant to flying skills. The nature of the laboratory study of retention permits, for the most part, the promotion of only initial learning in the individual, which again is quite irrelevant to describing the effects of disuse on reasonably experienced pilots. Similarly, interim periods used in the laboratory are most often hours and days rather than the longer intervals of inactivity hypothesized to affect flying skills (both normal and emergency reactions), and extrapolating from this is hazardous. Also, comparisons made among studies yield the disquieting feature that the considerable dissimilarities in the methods used prevent equivalence in comparing results (e.g., many different tasks, variations in original learning, different measures of retention). It is, of course, obvious that the study of retention under controlled conditions is a most difficult undertaking because of the number of major parameters which affect performance and because of the necessity for installing sufficient time intervals between training and test of retention. Naylor and Briggs (1961, p. 2) say, "The number of problems surrounding retention research increases rapidly as the duration of the retention interval is lengthened. How, for example, can one quantify and/or control the interim activity when its length extends into days, months, and years? How does one arrange for subjects who would be available for repeated testings over long periods of time? These and other considerations have contributed to the reluctance (and perhaps inability in terms of facilities) of researchers to attack the question of long-term retention with more vigor."

A number of generalities have emerged from these studies. An indication of these is set forth here to acquaint the reader with the intensity of the findings. Following this, the motors skills research most pertinent to pilot training, and studies of retention using flight simulators, will be summarized.

A consistent finding of the studies concerned with retention of learned behaviors is that greatest forgetting occurs immediately after

cessation of practice, decreasing in magnitude over time, which Ebbinghaus expressed as a negatively accelerated logarithmic curve (see Katona, 1940). Forgetting is retarded when the activity is overlearned, i.e., retention is highly correlated (showing a negatively accelerated curve) with the amount of original learning. The most rapid forgetting occurs with learned verbal material and learned procedures especially if the material lacks associative structure. Some data indicate that learned concepts are somewhat resistant to forgetting and that continuous motor skills are most successfully retained primarily because of their organization (i.e., task integration). A number of studies indicate this and are cited in various sources. For example, see Woodworth and Schlosberg (1954); Hovland (1951).

Qualitative changes also occur with the passage of time as evidenced by 19th century studies of changes occurring in memory for figures (see Woodworth & Schlosberg, 1954). More recently, research has been devoted to understanding the regulatory tendencies in people that presumably are due to cognitive consistencies in perceiving and remembering. Several cognitive control dimensions have been identified to account for qualitative differences in remembering. Among these are: leveling-sharpening, which refers to individual differences in degree of differentiation of memory traces (i.e., regularizing or accenting certain parts); focusing or scanning, which refers to differences in extent of attention deployment; and equivalence range, which reflects the preference of people in categorizing perceived similarities and differences. (See Messick, 1961, for an accounting of these and other cognitive control variables).

Retention of Perceptual-Motor Skills: The motor skill studies have emphasized the feature that the amount and extent of forgetting is related to task characteristics, the case in point being that uniformly greater retention accrues in performing continuous, coherent-sequence motor tasks than in performing discrete tasks. Melton (1964) investigated forgetting in a tracking task as a function of task characteristics, the hypothesis being that retention varies to the extent that display-control relationships are compatible with population stereotypes. The Battelle Electronic Tracking Apparatus (BETA) provided the task for 28 subjects in each of 12 groups (2 x 2 x 3 factorial design). The factors were: random vs. nonrandom target pattern; normal vs. reverse display-control relationships; and retention intervals of five minutes, one day, and one week following 20 trials given on the first day. Some results indicated that the normal (compatible) display-control relationship was significantly superior for retention than was the reverse (incompatible) relationship. There was no evidence that the differences in performance between the compatible and noncompatible modes decreased

even with extensive practice. The beneficial effect of practice was shown for both modes when the target motion was nonrandom (patterned, coherent), but not when it was random. Thus, the hypothesis appeared to hold for coherent but not for random target motion (i.e., greater retention when the task had greater organization). The evidence from this study, however, is not clear, and at best, suggests guidelines for further research.

A number of laboratory investigations have shown that continuous-control perceptual-motor skills are well retained over fairly long periods of no practice. Typical of these is a study by Fleishman and Parker (1962) on skill retention in a complex tracking task which simulated characteristics and control requirements of an airborne radar intercept flight. Seven groups of 10 subjects each were retrained following no-practice intervals of from 1 to 24 months. Correlations in the .80s and .90s were obtained between terminal training performance and retention performance for the groups. For groups initially trained to high levels of proficiency, virtually no loss was observed for periods up to 14 months of no practice, and the losses that did occur were recovered in the first few minutes of relearning. After 24 months of no practice, recovery occurred within the first 20 minutes of relearning. Variations in retention interval (1 to 14 months) were unrelated to retention performance. The level of proficiency of trainees was the most important factor and was just as important for short and long periods of no practice. Trainees having a high level of original learning retained more than those having intermediate levels, and this relationship of rank held throughout.

In retraining pilots it is important to know whether primary attention should be given to flight control tasks or to discrete procedural activities. Ammons et al. (1958) investigated this problem using both a complex tracking task (airplane control test) and a procedural task requiring sequential control manipulation. For periods up to two years, much better retention was obtained with the continuous task. A variety of other studies showing similar trends have been reviewed and rereviewed (see, for example, Adams, 1961; and Naylor & Briggs, 1961). The consensus on the topic of retention and task organization (see, for example, Naylor & Briggs, 1961) is that there is no evidence that motor tasks are intrinsically less susceptible to forgetting than are other kinds of tasks. A motor task that does not possess a meaningful patterning of responses (i.e., responses in random order or lacking in logical sequences of motor adjustments) may involve rapid forgetting with disuse.

From the point of view of the psychologist attempting to understand mechanisms of human behavior, this hypothesis may be entirely correct.

However, from the point of view of one interested in the retraining of aviation personnel, the evidence is clear-cut that flight control skills are well retained over quite long periods of time, while procedural activities are rapidly forgotten.

Evidence suggests that the decrements of massing are important to performance but not to learning, and that the residual effect of distributed vs. massed practice on subsequent performance is slight. Massing produces a decrement in performance which may persist throughout the practice periods but this does not necessarily mean that the individual has learned less. A test of retention may show greater learning than has been indicated in practice (Reynolds & Bilodeau, 1952; Gagne, 1953). For more complex tasks, whole learning may yield less forgetting.

Fleishman and Parker (1962) studied the effectiveness of different retraining programs. It was found that during initial retraining, mass practice versus distributed practice produced virtually identical improvement. However, in the fourth retraining session, as shown in Figure 10,

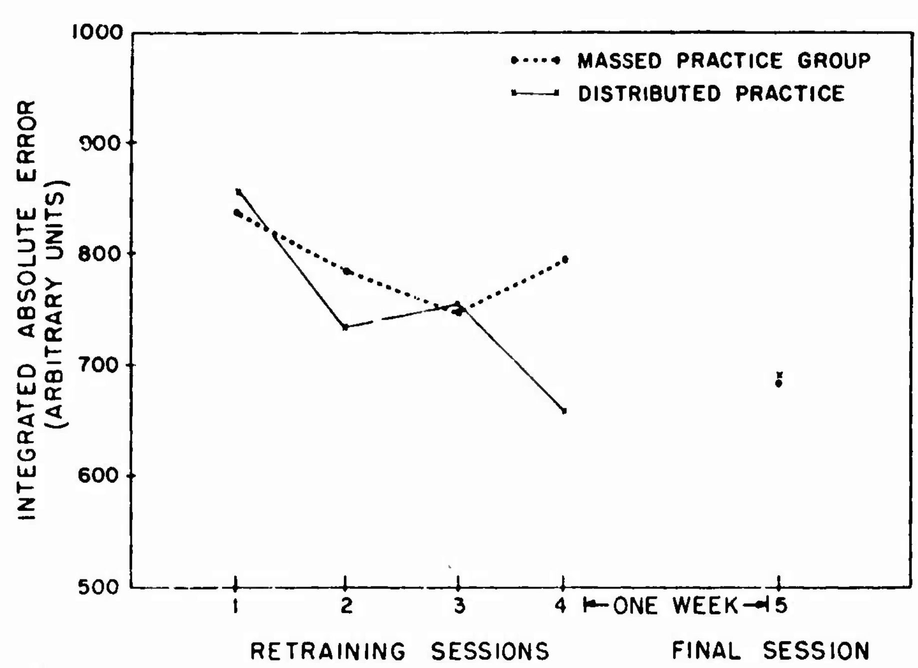


Figure 10. Effect of Different Retraining Programs During Retraining and After a Further One-Week Rest (from Fleishman & Parker, 1962.)

the distributed practice group was found to be superior to the massed practice group. This difference, which is statistically significant, may be due to some cumulative effect building within the massed practice group. Of particular interest, however, is the fact that when measured one week following the final retraining session, the performance of the two groups was virtually identical. Massed versus distributed practice apparently is of consequence during the retraining sessions, but is of no consequence in terms of transfer to later performance.

The value of rehearsal in increasing the retention of motor skills during periods of no practice has come under scrutiny. Naylor and Briggs (1961) summarize studies which indicate the value of interpolated verbal behavior, conducted during no-practice periods, for increasing retention (e.g., reconstruct the task mentally or on paper, think through the whole sequence, etc). They conclude from the literature that, to be most beneficial, the rehearsal task must have high fidelity to the originally learned task; also the more overt the rehearsal activity, the greater the facilitation.

Habit Interference: An important factor in performance decrement is the effect of intervening learning on the retention of earlier learning. The learning of new material or activities may not only impair the retention of material or skills learned earlier, but may also give rise to confusion when the individual is learning a new mode of responding while still maintaining proficiency on an earlier learned task. This habit interference is particularly noticeable when tasks are poorly designed, e.g., violating a population stereotype. Habit interference may occur in situations where a pilot is required to learn new responses to stimuli connected previously with other habitual responses. It may also occur when newer and somewhat related responses are interposed with older learned responses (e.g., alternately flying two different types of aircraft). Often, a "temporary disruption" in behavior-type forgetting occurs which may result in disaster in the air. A large volume of literature exists on interference effects in performance involving the considerable data on retroaction and transfer of training. The relationships are too lengthy and complicated to warrant a discussion here. A resume of this literature on motor habit interference is provided by Smode, Beam, and Dunlap (1959).

Simulator Performance: Two studies involving the use of a flight simulator confirmed the general literature finding that greater retention occurred in continuous skills than in procedural (discrete) responses. Mengelkoch, Adams, and Gainer (1958) studied the effects of a four-month interval of no practice on instrument flying skills. Following

four hours of ground training and one familiarization trial in the SNJ Operational Flight Trainer (Link 1-CA-2), two groups each of 13 ROTC students (naive to flying) were given five training trials and ten training trials respectively (defined as "intermediate" and "high" amounts of training). After a four-month interim period, the two groups were given four retention trials which were identical to the training trials. Each trial was a 50-minute instrument flight involving a series of maneuvers and procedures. Flight control was scored in error deviation from an index of desired performance recorded each ten seconds for airspeed, altitude, bank, rollout on heading, and leveloff at altitude. One hundred and twenty-five procedural items were scored on an error/no error basis (wrong, omitted, or out of sequence). In addition, two emergencies (propeller overspeed, and fuel warning) were given and measures taken of altitude lost, minimum airspeed achieved, and time to complete the sequence. The conclusions were similar to those in earlier studies, namely, discrete responses were more susceptible to forgetting than were continuous tracking responses. All measures showed a substantial retention loss. For example, the mean percent retention loss for the continuous-control flight parameters showed an excess of 30 percent when all parameters were combined for both groups (more than for the procedural groups). It was stated, however, that the percent loss values for the intermediate group indicates a loss in much higher proportion to the amount learned in training due to an artifact in the computational formula. The authors argue that the absolute amounts of losses, even where statistically significant, have no important operational meaning. For example, the mean loss of altitude for both groups was about 19 feet and between 0.9 and 2.4 MPH for airspeed. The loss of between 16 percent and 20 percent in procedures for the two groups, however, suggests a more serious degradation for flying proficiency. Thus, greater emphasis on procedural training is the sensible conclusion for the training of pilots. This conclusion notwithstanding, one wonders if five and ten trials respectively in an SNJ simulator is sufficient to warrant the venturing of a rejoinder for the training of Air Force pilots and if the tasks used in the SNJ device relate meaningfully to continuous control and procedural requirements of complex air weapons systems.

Another study (Hufford & Adams, 1961) investigated the effectiveness of a cockpit procedures trainer (CPT) for relearning aircraft procedures forgotten over a 10-month period. Twenty private pilots who had participated in an earlier experiment on the use of the CPT in original learning of procedures (Adams, Hufford, & Dunlop, 1960, experiment described on p. 164 of this report) were used. Ten subjects belonged to the original experimental group and had used the CPT for learning procedures. The other ten subjects were in the control group and had only whole-task practice throughout original training. The

criterion task (whole-task procedural and control responses) was a hypothetical and simple toss-bomb maneuver three minutes in duration. The control group was given ten whole-task trials (SNJ simulator, i.e., Link 1-CA-2) flown from memory, with knowledge of results given after each trial. The experimental group was given ten CPT trials flown from memory, also with knowledge of results given after each trial. Each subject was given ten whole-task trials just like the control group. The results revealed almost complete forgetting of procedures (95 percent correct at the start of the retention interval; 5 percent correct ten months later). CPT practice significantly enhanced procedural performance in the criterion task; however, the level of whole-task performance in procedures after CPT was significantly below that held ten months earlier at the onset of the retention interval. The earlier performance level was reestablished after two trials of additional practice in the simulator. The authors contend that a CPT does not suffice for complete restoration of procedures that must be performed in a time-shared relationship with flight control activities. Timesharing is a factor for the recall and relearning of forgotten procedures just as it is for original learning. As in original learning, some additional whole-task practice (in this case, two trials) is required before proficiency is maximized. These limited data suggest that while the CPT is a useful training device, it must be used along with the simulator, since some integrative whole-task practice must follow procedural sequences learned in the CPT. Thus, this represents a duplication of training function since all training can be accomplished in the simulator. The CPT will, however, reduce simulator utilization time.

What essentially emerges from the studies dealing with retention of skill? At best, they provide only guidelines for the development of research in the aviation environment. The use of tasks quite dissimilar to flying, the involvement in only the barest initial learning, the use of naive subjects, etc., all contribute to the irrelevance of, or the lack of generalizability of, the results for solving operational problems concerning the effects of disuse on pilot proficiency.

Research Issues: Research on retention, as presently conceived, has reached an impasse. The same conclusions continue to be generated with very little additional substance added. Entirely new strategies for research are required, set in more operational situations involving complex behaviors defining pilot performance. The obvious start is with programmatic research which systematically investigates the complex of variables subsumed under retention and utilizes various and lengthy time periods between learning and recall.

The most realistic effort requires the study of forgetting in terms of the characteristics of complex tasks. An understanding is needed of how behaviors are forgotten in complex tasks, and how the forgetting of the constituent components or elements of behavior is related to the total or gross activity. A desired output would be the determination of differential deterioration rates for the classes of tasks found in the pilot's job. For practical purposes, it would also provide information for the development of a "core" of activities and procedures to be installed to aid pilots in retarding forgetting, i.e., "staleness" resulting from periods of no flying. Information on the onset of decrement and the shape of the curve of forgetting over time for each task class should be an eventual goal of this research for pilot training. Retention studies over long periods of time are implied, since the effects of long-term retention of skills is so important in aviation. For example, infrequent emergencies call into play certain behaviors which have not been directly practiced for lengthy periods; yet these skills must be maintained at an adequate level to be employed when needed in the control of the aircraft.

The organization hypothesis that task coherency or integration (i.e., sequential-logical activities, nonrandom patterning, etc.) enhances retention is worthy of exploitation in terms of the pilot's job. If meaningful research is to be conducted, an effort is required in equating types of tasks in terms of level of difficulty.

In order to assess different amounts of forgetting, an understanding of the levels of learning for classes of tasks is required. Since types of tasks are measured differently, a common scale of measurement, and an understanding of the relations between measures is a necessary outcome. At present, the relationships between the criterion scores used in the many studies in the literature are not known, hence the difficulty in comparing results among studies.

Performance Decrement Over Time

A number of studies have attempted to determine whether there is a decline in the pilot's ability to maintain a consistently high level of flying performance under "fatiguing" conditions. Our review considers those which have employed objective measures of fatigue effects on performance, particularly in tasks involving pilot control of his aircraft over time. Also considered are vigilance decrement and work-rest cycle studies which are examined for their relevance and implications for pilot tasks. In reality, performance decrement is more truly associated with flights of longer duration than is flown by current aircraft. It is more a problem for advanced flight vehicles, having less meaning

for pilot training in our established frame of reference. No attempt has been made to review the literature on fatigue. The reader interested in the subjective aspects and general nature of fatigue is referred to a number of discussions (e.g., Bartley & Chute, 1947; Broadbent, 1958; Finan, Finan, & Hartson, 1949; and Mohler, 1965).

During World War II, Bartlett (1942) investigated the effects of fatigue on pilots' flying performance and found that pilots tended to show signs of fatigue after about 2 1/2 to 3 hours of flying. The demonstrated signs were: (1) control of steering column and rudder became less smooth and accurate over time and large errors (although made less often) took the place of smaller errors; (2) timing of coordinated movements became less accurate with rate of climb and dive, and control of direction being especially affected; (3) there was a tendency not to attend to those instruments infrequently monitored and those actions not regularly accomplished, and; (4) as the pilot tired, he was likely to do one thing without much reference to other related things. This was especially noticeable in instrument flying. Bartlett found that both experienced and inexperienced pilots were liable to fatigue and that the signs of flying fatigue were substantially the same from individual to individual. He noted that the pilot himself often did not know what the signs were or when they began to affect his behavior. He further noted (p. 3) that, "the signs of flying fatigue rarely, if ever, indicate a state in which the correct behavior or the desired skill, <u>cannot</u> be performed but only a state in which they <u>will</u> not be performed unless particular care is taken."

Davis (1948), reporting on the Cambridge Cockpit studies, found similar deterioration of various aspects of pilot performance over time. Especially prominent was the deterioration in control operation. McIntosh, Milton, and Cole (1952) performed airborne exploratory research to determine which, if any, pilot abilities deteriorated during extended instrument flights. Three pilots flew a C-47, one for 10 hours, one for 15 hours, and the third for 17 hours. Onboard equipment was used to record (1) the amount of time flight indicators were kept within tolerance limits and (2) the continuous variation of flight indicators and controls (17-hour flight only). Variables scored were airspeed, altitude, heading, vertical speed, bank, and pitch. The authors reported that pilots kept their flight indicators within specified tolerance limits for both precision maneuvers and straight and level flight as well after 10, 15, and 17 hours of instrument flight as they did during the first few hours of these flights. Comparisons were not made between flights. The continuous variation measures also gave indications that performance, as measured, was not a function of time since no decrement appeared between the first and last portions of the flight. There was a

suggestion however, that performance "may" deteriorate in the middle of an extended flight. The authors emphasized the exploratory nature of their research and the fact that the pilots knew the variables on which they were being scored and when scores (data samples) were being taken. Thus, an enhancement effect may have been present due to the increased motivation to perform.

Jackson (1956) attempted to obtain more typical, or standard, measures of pilot performance over time by not calling the pilot's attention to the variables being scored or the time of scoring. Measures of extent of deviation (rather than time-out-of-tolerance limits) from heading and altitude were taken. Using a more sophisticated design, he evaluated the relative performance of ten pilots comprising four air-crews, structured in such a way as to permit evaluation of changes during a two-hour watch, over a 15-hour flight (eight watches, generally four per pilot) and over four days of such flights. All measures were taken on a straight and level course. It was found that (1) performance in maintaining a constant heading deteriorated during 40 minutes of continuous work, (2) performance in both heading and altitude deteriorated during the first three of a pilot's watches and partially recovered in the fourth, (3) during the first two watches, pilots tended to fly more accurately and consistently in rough air than in calm air but that in the last two watches they were adversely affected by turbulence, (4) performance did not change appreciably from flight to flight during a week in which four 15-hour flights were flown on alternate nights, and (5) the deteriorations observed could not be accounted for by increased turbulence. Statistical evaluations were inconclusive.

In a recent investigation of performance deterioration over time, Hartman (1965) required four Military Airlift Command pilots to fly 24-hour simulated transport missions. Eleven successive legs were flown with each leg terminating in an ILS landing. Flights were flown in either a C-124 or C-133 simulator. Primary performance measures included data samples of airspeed, altitude, and rate of climb or heading through the cruise portion of each leg. The ILS ground track record displayed at the simulator instructor's station was photographed after each approach, and proficiency on ILS procedures was evaluated from these records. The results showed that, in general, performance (control operation) was maintained at a constant level throughout the inflight portion of the 24-hour flights; however, there was a trend in the data, although not significant, for systematically larger errors as the flights progressed. On the ILS approaches, proficiency was maintained at a constant level until 20 hours (tenth leg) when a substantial decrement occurred. Recovery from the decrement occurred on the final leg.

The author makes specific note of the motivating conditions surrounding the simulated flights and the psychological and operational flying support of copilots and flight engineers who "flew" with the pilots.

A component of flight that is especially susceptible to fatigue effects is low-altitude, high-speed flight. Simulator studies on LAHS flight conducted by North American Aviation (see p. 70) have found that fatigue effects do not become prominent for at least three hours of such flight. One study found no fatigue effects or physiological disturbances during a 90-minute flight. Others have shown that vertical acceleration in tracking does not produce sufficient fatigue to affect performance in three hours of a severe acceleration environment. Another indicated that three hours of continuous buffeting with severe vertical accelerations can be tolerated, although in this case the risk of incapacitating fatigue is high. It must be recognized, however, that in these simulator conditions the pilot is never in any danger and consequently the results of such studies are not completely satisfactory for generalizing to the actual flying environment.

One study that did investigate low-level (low-speed) flight contains a suggestion that actual flying performance does deteriorate over a relatively short time period under demanding flight conditions. Lewis (1961) investigated the performance of Canadian Army pilots required to navigate accurately over unfamiliar terrain at low altitudes (25-100 feet) in a 100 mph L-19 aircraft. Four Army pilots flew three, 2-hour, low-level sorties, each in quick succession on each of four days. In these exacting flights, there was a tendency for track-keeping performance to deteriorate in the second and third sorties of the day as determined by percent of time outside (arbitrarily chosen) flight corridors of 1/4, 1/2, and 1 mile widths. Frequency of excursions outside a 2-mile-wide corridor also increased during the second and third sorties as did the tendency to "fly into wires" and to "forget to switch fuel tanks." Errors in maintaining track were said to stem from (1) setting course in the wrong direction (rare), (2) searching for the end point of the flight, and (3) failing to make a correct course change for an "appreciable" distance after losing track.

The evidence for longitudinal performance decrement in present-day flying is contradictory. While there do appear to be certain changes in performance over time, there does not appear to be a loss of basic ability. It is conceivable that the changes that do occur reflect a lowering of the pilot's subjective standards for performance (seeCambridge Cockpit studies in Davis, 1948). As pilots become fatigued they may be less willing to exert the extra effort required to maintain their normal per-

formance standards. Jackson's results (1956) support this notion since he found that the relative performance of pilots was better under poor weather conditions in the early parts of a 15-hour flight, but late in the flight, performance was adversely affected by turbulence. McIntosh, Milton, and Cole (1952) similarly reported that pilots lowered their performance standards (were willing to tolerate more error) as a function of time in flight. The trends in Hartman's data (1965), although not significant, also suggest this phenomenon. Such findings have generally been interpreted to reflect motivational change over time rather than true performance decrement (cf. Ray, Martin, and Alluisi, 1961; Pearson, 1957). It is difficult, however, to interpret flight studies of dangerous low-level navigation in this light.

Studies of vehicle operators driving for relatively long periods of time were also examined for their relevance in resolving the question of decrement over time in operational control tasks. Most of these driving studies used inferential measures (vigilance, reaction time, etc.) rather than measures of control behavior over time and hence were excluded from consideration. One study, however, did use actual driving behavior as the criterion task and is reviewed here. The performance of male, military vehicle operators was assessed on a driving test battery by Herbert and Jaynes (1963). Subjects (independent groups at each hour point) drove a 3/4-ton truck for one, three, seven, or nine hours over a "fatigue" course. Following the requisite number of hours of driving, subjects performed a series of driving maneuvers which included moving the truck up or down a 15° slope, moving it about in a stall enclosure, parallel parking, etc. The results of the study demonstrated that there was a progressive deterioration in driving skill up to (and including) the seven-hour point. Subjects in the nine-hour group performed better on the driving test battery than the seven-hour group. The meaning of this last improvement was not apparent.

In summary, it is not at all clear if the pilot's control of his aircraft is affected by time spent working. Decrements have been shown and these may be of serious consequence despite the fact that they can be attributed to the pilot's willingness to accept more error in his performance over time rather than to real losses of proficiency. Conclusions that the pilot could, given appropriate motivation, produce behavior indistinguishable from that of well-rested subjects at any time that he wished, must be weighed against the number of aircraft accidents attributed to "fatigue" (Mohler, 1965) and to the stresses of demanding flight regimes such as low-level flying. Whether such stresses lead to rapid fatigue buildup, and in turn, to performance deterioration is unknown. Studies which imprefectly duplicate the extremes of temperature,

fear, anxiety, vibration, concentration, alertness, etc., inherent in the flying environment (e.g., simulator and laboratory studies) will be limited in providing definitive answers. In-the-air assessments have been limited by lack of adequate measurement techniques, differences in scores and tolerance limits, procedural differences, and nonstatistical treatment of data. Studies of vehicle operators suggest that long hours of driving affect performance adversely on measures taken after the driving experience but unfortunately shed little light on how performance changes during the couse of driving. Admittedly, such data are there if one is willing to make the additional inferential step that, for example, behavior of a group assessed after three hours is representative of the three-hour point performance of a group driving for twelve hours.

Vigilance Decrement: Much laboratory research on human monitoring behavior has been conducted in the laboratory under the rubric of vigilance. Summaries may be found in Bergum and Klein (1961); Buckner and McGrath (1963); Jerison and Pickett (1963); McGrath, Harabedian and Buckner (1959). Unfortunately, the relevance of much of this laboratory research to pilot monitoring tasks is difficult to establish clearly. Laboratory research usually requires a passive observer to respond to the presence or absence of weak, brief, infrequently occurring signals. Simple motor acts are typical responses and the interpretation of signals is not usually required. Few areospace monitoring tasks fit this description. In the flight environment the pilot is actively engaged in other tasks and monitoring is timeshared with these other requirements. Signals, well above threshold, usually persisting for some period of time, arise from multiple information sources. Kibler (1965) has noted that technological change has reduced the number of tasks which have characteristics approximating those typically employed in laboratory vigilance research; and field studies of radar and sonar operators (Elliott, 1960) suggest that performance decrement as identified and studied in simple vigilance tasks may not be a major problem in modern monitoring tasks. Although no direct assessments of pilots' monitoring performance over time could be found in the literature, the available evidence suggests that performance decrement over time may not occur when subjects are engaged in active tasks (Adams & Chiles, 1960; Dobbins, Tiedemann, & Skordahl, 1961; Mast & Heimstra, 1964) and when subjects are monitoring multiple information sources (Dobbins, Tiedemann, & Skordahl, 1963; Montague, Webber, & Adams, 1965). There is no evidence that vigilance decrement is a major problem for pilots in the present mission context.

Work-Rest Studies: Studies pertaining to the scheduling of work-rest conditions (e.g., Adams & Chiles, 1960; Alluisi, Chiles, & Hall, 1964; Ray, Martin, & Alluisi, 1961) for maintaining a given or desirable level of performance over long durations were also considered for possible inclusion in the report. While these studies pose important considerations for extended spaceflight and for the efficient utilization of personnel over long time periods, the implications for pilot training are limited. Consequently, no review has been attempted.

Research Issues: The extent to which the pilot's ability to control his aircraft changes with time is unknown. The literature suggests that while performance does worsen, there is no loss of basic ability. This topic should continue to be studied but within a realistic time frame for piloting tasks and within the airborne environment, using instrumented aircraft to determine the magnitude and consequences of change over time. Simulator studies are less useful for making such assessments.

Another area for research concerns the interaction of stress with fatigue. Pilots flying low-level stressful missions may reach a performance limit in a shorter time period than pilots flying more routine missions. Some consideration should be given to determining safe limits for number and duration of such flights.

Bartlett (1942) has noted that the pilot is usually unaware that his performance is changing over time. It is conceivable that training routines can be developed to assist the pilot in maintaining a required state of alertness when flying long-duration missions, and some effort should be expended in this direction.

SECTION IV

STUDIES ON IMPROVING THE TRAINING SYSTEM

Studies pertinent to pilot training have been conducted which are not easily placed in the preceding sections. These can most reasonably be defined as (1) investigations yielding information for improving training systems in general and (2) studies of ancillary aspects of pilot training which, although important, have not received primary emphasis in training programs. The latter represent areas which may be considered on the "fringe" of training, and are supported by minimal empirical evidence. Thus, a collection of studies heterogeneous as to content and purpose are assembled in this section to call attention to the status of these areas insofar as pilot training needs are concerned.

Emphasis in the discussion is placed on method or technique (or on the status or condition of a researchable area) since the research appraised is, in most cases, exploratory and a body of definitive data that has implications for pilot training has not yet coalesced. Three topical areas are identified into which the studies in this section are grouped. The first group of studies concerns the development of methodology for determining training requirements. These researches center on the development of methods for improving the derivation of this necessary information as it applies to any large training system. Since pilot training systems represent a specific case of training systems in general, the concepts and procedures developed in this body of research are highly appropriate to the purposes of this report.

The second group of studies describes those recent innovations in training method and technique that are applicable to pilot training.

The final group of studies considers a number of issues specific to the aviation environment that are not given primary consideration in the training of pilots. These studies deal with diverse issues in flying training, and are not easily cataloged as a group. They are brought together here under the innocuous rubric of "additional research areas."

Accordingly, the discussion in this final section follows this outline:

<u>Development of Training Requirements</u>

Development of Training Objectives
Task Classification
Performance Standards

Innovations in Training Method

Adaptive Simulation of Vehicles
Computer-Based Instructional Systems
Self-Confrontation Techniques

Additional Research Areas

The Flight Instructor
Physiological Indoctrination
Escape Training
Sensory Training
Attitudes Toward Equipment

DEVELOPMENT OF TRAINING REQUIREMENTS

The present decade has been witness to an expanding technology of training characterized by an orderly development of knowledge and method applicable to the design of training systems. Perhaps the key phrase underlying this development is that it is a "systems approach to training." This approach is defined by an organized instructional program with precise goals and defined interrelations between system components. In its fundamental terms, the design of a training system can be described as follows. The behavior which must be exhibited on the job is the goal of the training system. To achieve this, an integrated series of learning experiences must be provided, organized in a defined time frame to produce the required behaviors. Testing is repeatedly conducted to assure that training design is compatible with job requirements. The process begins with the defining of training objectives (e.g., desired performance outcomes) which provide the basis for deriving training content and for developing criterion measures required to test the training system. The design of training methods, training hardware, and materials are based, in turn, on training content. These methods and materials, when administratively implemented, describe the training program. Criterion measures provide indications of the adequacy of the program outputs, and feedback loops determine the adequacy of the system output. A simple schematic of this systems approach to training is shown in Figure 11 below. These components and their interactions have been articulated in the writings of such leading scientists as Eckstrand (1964), Crawford (1962), and Gagne (1962).

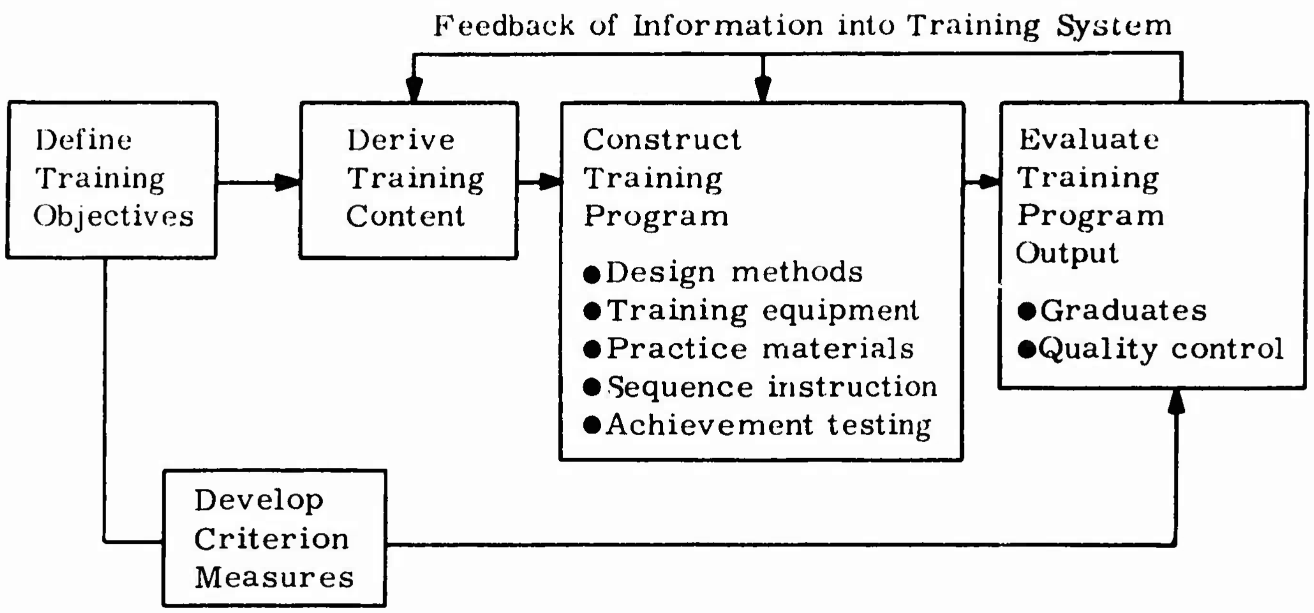


Figure 11. Schematic of the Systems Approach to Training

The activities and requirements involved in the systematic development of training programs can be subsumed under three major elements:

1. Determining the requirements for training.
2. Developing an environment for the conduct of training.
3. Measuring the progress and outcomes of training and controlling the quality of student output.

Our interest in this section centers only on the researches dealing with item (1) above, i.e., studies which develop concepts and methods for specifying training requirements pertinent to pilot training programs. Aspects of items (2) and (3) above have been treated elsewhere in this study. This area reviews contributions to improving training systems. The discussion appraises the studies on methods for developing training objectives; reviews studies on job knowledge, skill determination, and task taxonomy; and considers research on development of performance standards.

Development of Training Objectives

An important contribution for improving the design of training systems has been the emphasis on precision in specifying training objectives and the development of methods for accomplishing this requirement. Well-defined, job-relevant objectives provide the means for an orderly

development of training content. The basis for this development is that objectives, to be meaningful, must be specified in terms of what the trainee is required to know in order to perform the job. Certainly, not all training programs have been installed on this basis. The philosophy of learning a job while on the job and matching job titles with textbook and subject matter titles is quite different from that of designing training for a set of job-tasks in terms of the performance expected of a trainee at course end. Miller (1962) states that many training efforts include fads and local prejudices, including what seem to be "initiatory and tribal rituals." Thus, the emphasis on training objectives in terms of job-performance expectations, serves in differentiating the irrelevant and ritualistic from the essential.

The studies in the literature dealing with training objectives emphasize method and provide means for the systematic development of job-relevant training objectives. Discussion is also devoted to various sequences in the process of developing training objectives as the basis for efficient training system design. While differences exist regarding nomenclature and time at which an activity is performed, a core of general steps is identified below which summarizes the current thinking on training objectives development (Eckstrand, 1964; Smith, 1964). The process begins with an analysis of job-required behaviors.

1. System Definition. The initial step is to understand the system to which trainees will be assigned after graduation. This provides an overview of the job context and enables the activities of the trainee to be related to the mission of the system.

2. Task Inventory. This is a listing of all duties and tasks in a job. From this, the tasks to be trained are identified, i.e., analyzed to determine the difficult, critical, trainable tasks. Thus, specification of what tasks should be taught and, for these, what level of proficiency should be required, are two outcomes of this effort.

3. Task Descriptions. These are developed from statements of tasks to be trained. Each task is partitioned into activities described in operational terms in sufficient detail to provide a basis for the development of skill and knowledge requirements. This description organizes the cues, actions, and indications of action in each task. The general format includes the nature of the activity or element, the equipment involved, the conditions underlying the performance of an activity, and criteria which define performance in terms of various measure classes. The task description forms the basis for the specification of knowledge and skill requirements of the task.

4. Knowledge and Skill Components. This is the process of determining precisely how the job performance requirements can be achieved. At present, the techniques of task analysis are employed to aid in the determination of what and how the trainee should be taught to perform effectively on the job. (See pp. 35-40.) A number of procedures and formats are available which attempt to specify the behavioral requirements in task performance. The adequacy of the training objectives is in large part dependent upon the completeness and accuracy of the specified skill and knowledge requirements. The process today is highly subjective and requires considerable experience of the training specialist. One attempt to improve this situation is the work on task taxonomy (discussed later in this section). Ultimately, a taxonomy is needed which will provide concepts and nomenclature for analyzing tasks in behavioral terms.

5. Statements of Objectives. The foregoing sequences are the prelude to the specification of training objectives which describe precisely what the trainee should do upon completion of training. Smith (1964) states that the objective should describe: the performance expected of the trainee, the conditions under which performance will be measured, and the standard (speed and/or accuracy) to be achieved. Eckstrand (1964) provides three criteria for assessing the adequacy of training objectives: relevance (knowledge and skill requirements adequate for job performance), completeness (all required performance outputs accounted for), and measurability (stated in a way to permit determination if objectives are achieved).

A precise knowledge of the behaviors to be trained is the point of departure for the efficient design and evaluation of a training system. Unfortunately, specifying the behavioral requirements in job performance is a difficult undertaking. The process of developing training objectives in terms of what the trainee must know to perform successfully on the job is complex and difficult and does not have an optimum procedure. As a consequence, training requirements are not systematically developed. The results of a recent survey of Air Force advanced pilot training programs (Smode & Meyer, 1966; Smode, Post, & Meyer, 1966) indicated that training objectives are specified in terms of judgment and experience from previous and similar systems and in terms of available or committed resources. Pilot training programs can benefit from strong development of methods for determining training requirements within the philosophy described above. The lack in methodology for determining training requirements, described above, is however, not peculiar to pilot programs; rather, it describes the status of training technology in this area. Thus a prime research requirement is the development of methods for determining training requirements totally relevant to job demands.

Task Classification

Experience with the design of military training programs has shown that the process of specifying training requirements is aided if job performance is viewed in behavioral rather than operational terms. Ideally, this involves grouping tasks in such a way that a correlative set of training requirements can be stated, the implication being that patterns of task attributes suggest specific training practices and policies. A task taxonomy which provides categories of behavioral elements, grouped to provide sets of distinct task characteristics, would put precision into task descriptions and enable more objective decisions to be made on what and how the trainee should be taught in order to do the job. The important feature of such task categories are that they are boundary-definable and independent of each other with respect to the procedures and media for promoting learning. The implications of this for pilot training programs and for training technology, in general, are obvious.

A search of the literature indicated considerable recent activity devoted to concepts and methods of task classification, and a number of studies have attempted to categorize performance for training design purposes. Individual contributions for military systems have been made by a number of investigators, notably: Fitts, Cotterman, R.B. Miller, Gagne, Hoehn, Glaser, Fleishman, Folley, Demaree, Smode, Stolurow, Willis, E.E. Miller, and Lumsdaine. The extent of this development is too detailed for review here. Fortunately, task analysis methodology and task taxonomies have been reviewed and evaluated for various purposes in papers by Eckstrand (1964), B.J. Smith (1965), R.G. Smith (1964), Chenzoff (1964), and R.B. Miller (1962).

Although there is the suggestion of much research effort, the best that can be said of the research to date is that it is exploratory and represents only initial attempts at developing task classification expertise. No standard scheme for classifying tasks or agreed-upon groupings of tasks exists today. Various writers have proposed anywhere from a few to more than twenty independent categories of behavior with varying sophistication in theoretical bases and richness in description. Yet, current taxonomies tend to be very similar in their nomenclature and coverage of behavioral requirements. For example, in operator tasks, the accepted words and phrases revolve around: procedural (fixed and variable) activity, perceptual-discriminative activity (searching, scanning, monitoring, discriminating, etc.), perceptual-motor activity (discrete, continuous), decision-making activity (concept using, problem solving, etc.), and crew activities.

Thus, the research has produced a number of, but slightly differing, catalogs of behavioral classes. These have no real utility since the superiority of any technique cannot be determined due to the absence of empirical data. The taxonomies represent, for the most part, opinions based on sketchy experimental data and high-level expertise about human performance in complex systems. They nevertheless, provide contributions to training research in that they represent a way to go in this difficult area. The available taxonomies should be the impetus for a substantial effort to test and evaluate their accuracy, reliability, validity, and heuristic power to determine if this is the right research course to pursue or whether new approaches and concepts are needed. The consensus is that the science of behavior in instructional situations will suffer until the problems of classifying the tasks on which instruction is given are solved (Eckstrand, 1964). It seems appropriate that the needed research should have recourse to theoretical systems relating task attributes to training variables.

A research issue for pilot training (also relevant to Section II of this report) is a detailed classification of pilot tasks for the purpose of organizing operational information in a way that would facilitate the application of principles, procedures, and media for the most efficient learning. The point of departure could be the identification and selection of common job activities in Air Force aircraft/missions combinations, based on such criteria as: frequency of occurrence of task elements, criticalness to system effectiveness, and amenability to training. A highly desirable result would be the identification of areas yielding greatest payoff in system effectiveness and the suggestion of training principles most effective for given task characteristics.

Performance Standards

The performance standard is a statement or measure of performance level that the individual or group must achieve for success in a system function or task. It provides information for specifying the levels of skill which must be developed for system effectiveness. Since it is based on performance requirements information (statements or measures which refer to system operation) the output should be the accurate and reliable specification of personnel performance on functions and tasks within a system, and the effects of variations in human performance on the system.

Job standards information is needed in the determination of training objectives and in curriculum development. The effectiveness of training standards is also dependent on the precision and validity of these standards. The training standard is an interim step between course learning and

performance on the job. Quite frequently, in training programs the standards for training are not compatible with performance requirements on the job, one reason being the lack in method for stating performance requirements objectively and precisely. Performance standards today are established on the basis of experience and expertise. While these judgments may be adequate for many systems, standards are rarely set correctly this way. When set too low, the result is degraded system effectiveness; when set too high, costly overtraining is the result.

An example of the type of work needed for developing objective methodology for determining performance standards is that being conducted by Dunlap and Associates, Inc. (1964, 1965), for the Navy. The research program is attempting to devise methods for establishing system-related job standards for Navy rates, such as found in major electronics systems (e.g., sonar, radar). The approach, which is mathematical, develops a set of personnel goals or minimum job standards with definable relations to system effectiveness requirements, in order to provide a basis for relating system performance requirements to levels of personnel performance. The job standard ultimately will be composed of: (1) a personnel/equipment functional unit, (2) an accuracy/time requirement, and (3) a required probability of successful performance. One output of this program has been a review of methods for evaluating personnel performance and general tools of system analysis (Hoisman and Daitch, 1964). No existing modeling or analytical techniques were found to be fully applicable to developing job standards. The Dunlap program, which is presently being pursued, is the most ambitious effort on job performance method to date and may serve importantly in structuring future research in this area.

INNOVATIONS IN TRAINING METHOD

Several relatively new areas of research have been selected for review because of their specific importance to pilot training. These areas have not yet come under systematic or heavy study; hence, a body of empirical data is not available. In most instances, basic or organizing concepts are still in the process of being clearly defined. Each of these areas is considered here primarily as a research issue on which exploratory work is currently underway. The research cited most truly reflects initial study of new concepts and methods that hold promise for improving training in aspects of the pilot's job.

No consideration is given in this report to innovations in engineering design or new hardware that may provide advantage to pilot training. These contributions will be presented in a following volume dealing with

engineering developments and advances which are or may be applicable to pilot training.

Adaptive Simulation of Vehicles

A simulation technique involving automatic adjustment of problem difficulty has received attention from individuals engrossed in tracking research. In this technique, called adaptive or self-adjusting simulation, a standard measure of continuous control skills is maintained and is employed to adjust automatically the difficulty level of the control task. The simulator is set to produce a "standard" amount of error and then changes the system until that amount of error is present (in conventional simulators, the system is fixed and the error is measured). This automatic adjustment in difficulty level can be made to vary by modifying the system input or forcing function, by changing the information displayed, by modifying the control signal, or by changing the dynamic characteristics of the simulated vehicle. As operator performance increases, the task automatically becomes more difficult; as performance decreases, the task becomes easier. The measure of performance is the difficulty level of the task.

Interest in current concepts of adaptive simulation go back to the late 1950's. Birmingham in 1959[24] described a technique in which the operator's error was used to modify the forward gain of the integrators in a tracking system, and hence its rate of response. The smaller the operator's error became, the quicker the tracking system became until he was tracking as fast as he was able. Birmingham referred to this technique as a means for measuring "human operator bandwidth."

A technique developed by Kelley presented at an IRE International Congress on Human Factors in Electronics[25] and later published (1963), involves averaging a measurement of the subject's performance over a

---

[24] Birmingham, H.P. The Instantaneous Measurement of Human Bandwidth, paper presented at the Eighth Annual Conference on Manual Control, held at Dunlap and Associates, Inc., Stamford, Connecticut, 11-12 May 1959.

[25] Kelley, C.R. Self-Adjusting Vehicle Simulators, paper presented at the IRE International Congress on Human Factors in Electronics, Long Beach, Calif., 3-4 May 1962.

period of time rather than using an instantaneous value of error (as did Birmingham). The average score is used to adjust automatically certain parameters of the task. The scoring technique for use with continuously distributed error scores involves setting an error threshold which is an amount of error that will not produce any change in the system. Error in excess of the threshold results in changes which makes the system easier, while smaller errors cause the system to become more difficult. Scoring involves an analog computer solution to the following equation:

$$S = K\int_0^t (e_L - /e/)\, dt + S_{initial}$$

where: S = the trainee's score, the parameter which determines the system difficulty, a higher score representing better performance.

e = system error (however defined)

e_L = a constant error threshold or crossover point, determining the amount of error that must be present before S declines and the system becomes less difficult.

K = a constant which determines the rate at which changes in S take place for a given amount of error.

The technique permits the control task to be at an appropriate level of difficulty for the trainee. He progresses at his own pace since the task is modified in the direction of the criterion condition as skill increases. This should facilitate skill training if we assume that training effectiveness is reduced when tasks are too easy or too difficult for the trainee. The record of the varying difficulty of the control problems provides a sensitive preformance index. The control task can be made intrinsically more interesting simply by providing a display which immediately tells the trainee how well he is doing from moment to moment (the S score). This "challenge" may increase the motivation to perform.

Other papers given at the same IRE Conference[26] also discussed adaptive simulation involving variations on the basic technique. These

---

[26] Hudson, E.M. An Adaptive Tracking Simulator; Birmingham, H.P., Chernikoff, R., & Ziegler, P.N. The Design and Use of "Equalization" Teaching Machines. Papers presented at the IRE International Congress on Human Factors in Electronics, Long Beach, Calif., 3-4 May 1962.

techniques differed somewhat from Kelley's approach in that automatic adjustment of "quickening" coefficients was investigated, i.e., they changed the parameter of display quickening in proportion to an average or filtered error score. In Kelley's technique, the equation (above) holds error constant and adjusts system difficulty changes until the set amount of error is present. In the Birmingham and in the Hudson techniques, error is decreased and the simulation parameter that is adjusted changes in some fixed relation as a function of improved performance (see Kelley, 1963).

An important study for adaptive technology was Hudson's (1964) development and evaluation of an adaptive tracking simulator. The constructed device provided a difficult pursuit tracking task (criterion task) that represented a high-gain, third-order system. Also, adaptive parameters of two types were manipulable: those that controlled the amount of assistance given the subject as a function of his error, and those that simplified the system dynamics as a function of the amount of assistance. Three levels of task difficulty were defined, i.e., easy, moderate, and hard. An evaluation was made of the device to test the hypothesis that superior training will accrue when the trainee is guided through an adaptive regime, where the level of difficulty is always proportional to his instantaneous ability. Seventy-three subjects were each given ten hours of training on the tracking device. A 17-inch CRT was the display used, on which appeared a circle (0.4 inch in diameter) that moved in a quasi-random pattern, and a dot. Each subject was required to keep the dot centered in the circle (pursuit tracking using a side-stick control). The task varied in either the dynamic nature of the controlled third-order system (gain, order and stability controls) or in the level of difficulty during practice. Thirteen experimental groups (ranging from four to six per group) were thus defined. The results indicated that almost all groups trained adaptively evidenced greater transfer than those who practiced only on the test conditions themselves. The adaptive group who practiced at the optimum level of difficulty showed 300 percent greater transfer than the nonadaptive group who practiced only on the test conditions. A main finding of the study was that level of difficulty during practice was the only important parameter. Optimum learning took place when the practice task involved moderate or average level of difficulty regardless of the level of difficulty of the criterion task. There was some evidence that an adaptive test, given during the initial stage of training, could be a more accurate predictor of final standing than could any form of data from non-adaptive tests. Hudson concluded that level of training is largely a function of level of difficulty during practice. Other parameters (e.g., type of change in the transfer function) are of minor importance except as they affect the level of difficulty. There is also some interaction between the

level of difficulty of the criterion task and the optimum level of difficulty of the practice task. The important implications of these findings for learning theory "is not the superiority of adaptive training, as such, but rather, the definite relationship between level of difficulty during practice and level of achievement on the criterion tests."

Recently, Kelley[27] proposed to the Air Force (Manned Orbital Laboratory Program) to employ a self-adjusting tracking test as a measure of an astronaut's ability to maintain piloting skills under conditions of flight (e.g., weightlessness). Preliminary data were obtained on two subjects given extensive training (180 trials over nine days). Each subject alternated on five-minute runs on an adaptive and a nonadaptive version of the test. The results indicated that the adaptive version of the test was consistently more sensitive to interday differences in performance due to training than was the nonadaptive version. This suggested that an adaptive test would form a more sensitive instrument for measuring operator skill.

In summary, adaptive training appears well suited for training in vehicle control. It permits the trainee to keep performing at the threshold of his skill level at all times. Current fixed-base simulators, for example, do not automatically keep the trainee working at an appropriate level of task difficulty. Training with fixed-difficulty level is often too easy or too hard for the individual, and one can assume that this feature reduces training effectiveness. Lost training time, then, is of some consequence. Limited laboratory data has suggested to Kelley (1963) that usual training methods may waste more than 50 percent of training time required for building up a high skill level.

As mentioned earlier, self-adjustment techniques may automatically vary: system input or forcing function, information displayed, the control signal, or the dynamic characteristics of the simulated vehicle. These may be differentially effective in a given training situation. Correct selection of adaptive parameters should provide optimal challenge to the learner, and supposedly, provide such instructional benefits as increase in learning rate, intrinsic motivation, and higher final levels of proficiency.

The research issues are many and interesting. In general, the field is new and research is needed to establish the concepts and potential of the technique and to evaluate the possible variations that exist. Adaptive devices have applications not only for training but for selection and

---

[27] Kelley, C.R., Personal communication.

human factors design (i.e., the values derive from the effectiveness with which these devices measure control skill).

The results of adaptive research emphasize a known principle that is important for training. The principle is that face validity is not a necessary condition for efficient learning. For example, Hudson's control group was trained and tested on the identical system (perfect face validity), yet this group learned very little. However, an adaptively trained group that performed under widely different practice and test conditions displayed a great deal of transfer of training. At present, training of pilots is done on complex fixed-based simulators which possess considerable face validity resulting from an emphasis on engineering fidelity. It is known that such training is inefficient. These complex simulators have the potentiality of being used as adaptive trainers along with their other capabilities. Flight simulator design is moving decisively towards this concept (adaptive circuits, programmed sequences for training) and substantial systematic research is encouraged in this area.

Many specific research inquiries are suggested, dealing with measurement of skill, rate and level of acquisition, and retention of skill in the adaptive mode.

Computer-Based Instructional Systems

It can be expected that the use of computer-based instructional systems will increase with time. Many claims have been made for the digital computer's value for instruction, ranging from its capability for such feats as storing whole libraries of readily retrievable technical material to its capacity for instructing as many as 1000 students at a time (e.g., PLATO III, see Hickey & Newton, 1966). Perhaps the most appealing claim, in terms of training value, lies in its potential for providing adaptive training, and it is this concept which is of prime interest in this part of the report. In adaptive training, the characteristics of the material successively presented to the subject for mastery are varied as a function of his preceding response(s).

Research at the University of Illinois (Stolurow, 1965) is engaged in the design of an ideographic model of tutorial instruction which is fully adaptive. The SOCRATES (a System for Organizing Content to Review And Teach Educational Subjects) has been developed and the feasibility for providing individualized instruction is being explored. SOCRATES is designed to make pretutorial decisions about the way in which a student learns from stored information about the student (e.g., achievement and aptitude test scores, entry-level skills and knowledges, etc.), and about

the teaching objectives (content, topics), the desired terminal proficiency level, and the time available for learning. The result of the pretutorial decision is a teaching strategy for the individual student. During the instruction process, SOCRATES makes tutorial decisions based on analyses of student responses to determine the what, how, and when of presentation. To date, research using SOCRATES has been aimed at collecting some of the enormous amount of data that will be needed to make such systems practical for training. It is apparent that computer-based adaptive training systems of this nature will be handicapped until much more is known about human behavior, learning strategies, the laws governing the acquisition of knowledge by humans, and so forth, since such inputs are needed for complete programs for instructing the computer. Nevertheless, even with current knowledge limits, such systems have utility for research and enormous potential for training.

While it appears that adaptive training holds great promise for skill acquisition in some task situations (see pp. 191-195), there are other situations where it is not desirable. Several studies on perceptual training have demonstrated this finding. Swets and his associates (Swets, 1962; Swets et al., 1964) investigated computer-based adaptive training for the identification of nonverbal sounds to discover the extent to which principles of programmed instruction applied in perceptual training. Their data suggested that training which automatically adapts to the performance output of the trainee is inferior to conventional training procedures. Another study (Weisz & McElroy, 1964) required subjects to learn to identify unfamiliar geometric forms which varied in four dimensions. Again, adaptive training was found to be inferior to other training routines. Mirabella and Lamb (1966), also concerned with perceptual training, compared the effects of adaptive and nonadaptive training on target detection performance where the stimuli consisted of symbolic data displays. No evidence was found that adaptive training was more effective than traditional procedures. The authors concluded that the use of complex training procedures requiring the speed and adaptability of a digital computer was not justified.

Computer-based instructional systems have also been used to provide nonadaptive training in several verbal areas (e.g., Coulson, 1962) and for psychomotor task training (Englebart & Sorenson, 1965) with results similar to those given by any programmed instruction technique (see pp. 62-64). However, when computers are used nonadaptively, more consideration must be given to balancing costs against training benefits or gains to be realized from their use.

The extent to which computer-based instructional systems can be gainfully employed for pilot training is largely a matter of research. Given a well-written instructional and computer program, it is obvious that much of the verbal material now taught conventionally could be taught via a computer-based system in nonadaptive, ordinary, programmed-instruction ways. Whether an adaptive training system has potential for pilot training is not precisely known. One approach to adaptive pilot training, however, has been described by the Aviation Research Unit[28] (of the Human Resources Research Office of the George Washington University) supporting the Army Helicopter School at Fort Rucker, Alabama. The proposed use of a computer-based instructional system for training helicopter pilots would consist essentially of presenting a series of standard flight lessons in programmed instruction format to the trainees. Difficulty of the material presented will be determined by the skill level of the trainee and, as skill increases, he will be "branched" to harder problems. Whether this is an efficient and effective way of training pilots must be empirically determined. Its implications for proficiency measurement and assessment may be more important than its training potential.

A great deal of research effort is needed to determine the value of the adaptive training concept for training pilot skills. If a given flying task is of a given inherent level of difficulty, is there any appreciable advantage to be gained by reducing the complexity of the task for programming purposes and presenting it piecemeal to the trainee to require him successively to master it in a progressive part fashion? The digital computer's inherent capacity for allowing adaptive training is insufficient reason to train individuals adaptively; it must be demonstrated that adaptive training is valid and desirable.

Self-Confrontation Techniques

Confronting a subject with recorded samples of his own behavior in a given situation has been shown to be an effective means of altering his behavior. This technique, known as self-confrontation, appears useful for training the pilot to communicate effectively in a growing number of situations in which he must interact in a positive and attractive manner with other individuals who may be his counterpart in friendly nations.

---

[28] SDR, Synthetic Flight Training System, 21 March 1966, U.S. Army Aviation Human Research Unit, Fort Rucker, Alabama.

For example, in special air warfare operations, it is desirable for pilots to be "culturally tuned" in order to interact competently with indigenous forces and civil populations. There has been recent interest in this methodology within the Air Force. A program of research on training for culture-contact and interaction skills is currently being conducted at the Aerospace Medical Research Laboratories at Wright-Patterson Air Force Base, Ohio. The overall objective of this program is to develop improved methods for training American military advisers to work effectively with their counterparts in other societies in support of counter-insurgency (COIN) or pre-COIN operations. The development of the self-confrontation technique into an effective means of cross-cultural training (particularly with respect to the nonverbal aspects of a culture) is one of the primary goals of this task (Haines, 1964). At present, one study has been completed (Haines & Eachus, 1965) using self-confrontation to train cross-cultural interaction skills. Pretrained subjects were put through a role-playing session in which they performed "critical" cultural behaviors in interaction with a "foreign" military officer (who was a member of the experimental team). After the session, each subject was confronted with a videotape recording of his behavior during the session. While viewing the tape, he was critiqued on errors and accomplishments and was then required to replay the scenario. Changes in behavior from the first scene to the second were evaluated. The principle result was that self-confrontation was shown to be effective for the rapid training of interaction skills although it did not particularly favor the acquisition of nonverbal over verbal skills. No attempt was made to assess retention.

The utility of the self-confrontation technique for complex skill training is being further investigated by this research group (Eachus, 1965). Other plans are to establish the full course of acquisition and retention through use of the self-confrontation technique, to experimentally establish baseline data regarding self-confrontation, and to identify more thoroughly the parameters which are manipulable with the technique (Haines & Eachus, 1965). Examinations of the relationships between subject attitudes and self-confrontation, as well as ways of simplifying the procedure of training through self-confrontation so as to reduce the expenditure of time and money, are also planned (Haines, 1964).

The self-confrontation technique has been successfully employed in the teaching of foreign languages (Carroll, 1963); for teaching retarded children good table manners (Haines, 1964); and in the training of teachers (Lumsdaine & May, 1964). Stoller (1964) has also used it successfully in psychotherapy. His technique consists of recording group therapy sessions of individuals suffering from chronic mental disturbances and then showing the tape to patients individually. Stoller

reports that self-confrontation has resulted in marked improvements in physical appearance, verbal behavior, and use of rational thought by the patients.

Nielsen (1962) found that subjects are more amenable to criticism and advice with this technique than they are with other critique procedures. He reports, however, that the value of the confrontation seems to diminish with time, the longer the interval between training and confrontation, the less the value of the confrontation in effecting behavior change. Presumably, this is because of the decreasing efficiency over time of the method as a means of simulating the recall of the various stimulus and behavioral elements of the original performance situation.

Self-confrontation appears to have value in teaching cross-cultural interaction skills to pilots assigned to military aid missions and in other situations requiring interaction with indigenous units. It may also be useful for initial pilot skill acquisition. For example, videotaping student pilots' early "flights" in simulators and then pointing out errors in a replay session may provide for more rapid elimination of these errors and quicker acquisition of skill. Crew training and other tasks requiring interaction skills may be similarly enhanced. Teaching communication skills using language laboratory methods should be considered.

Research should continue to develop and refine the method and determine its feasibility for pilot training situations. For example, continuing effort is needed to more thoroughly identify behavioral parameters that can be manipulated with this technique. The permanency of behavioral changes produced by self-confrontation is another issue worthy of consideration. Retention of skills learned from this method should be assessed to determine if it produces only temporary modifications in performance to fit the needs of a somewhat unique and, perhaps, sensitive situation for the trainee.

ADDITIONAL RESEARCH AREAS

This final grouping of studies deals with a number of diverse issues of "secondary" interest to pilot training programs. These issues are not easily placed in a framework for discussion, yet they are of importance to the efficiency and validity of pilot training. They are evaluated as a group under this imprecise heading to round out the content of this report. In most instances, few aviation-oriented studies are available in these areas, each area being better described as possessing problems that should be accounted for in pilot training. Because of this, the appraisal is brief.

The Flight Instructor

The quality of pilot training is in large part dependent upon individual instructor pilots. Despite the instructor's key position in flying training, little effort has been devoted to controlling the quality of instructor personnel, and to ways for maximizing their utility in training. A viewpoint that has prevailed is that since instructors are easily defined as expert pilots, their activities and procedures in instructing students are satisfactory to the objectives of the training program. Yet significant variability among instructor personnel in technique, philosophy of instruction, and performance assessment has been demonstrated repeatedly. One result has been a significant lack of control of their outputs in a training program.

Because of these considerations, the subject of the flight instructor has been included in this report, although it is not treated in detail. The research selected for review is concerned with the instructor as part of the pilot training process, specifically those studies which demonstrate that the instructor in some way influences the proficiency of individuals under his tutelage.

Variability Among Instructors: A study by Williams and Flexman (1949), previously cited, demonstrated that instructors differed significantly in judgment of when pilot trainees were ready to solo. These differences were sufficient to preclude an assessment of the effects of other variables (i.e., amount of simulator time) that were being investigated. Similarly, instructor variability regarding proficient flying performance has been a serious problem in assessing pilot proficiency and maintaining control over the quality of the training process (see pp. 121-123). While periodic attempts have been made to reduce instructor variability by training (Townsend, Flexman, & Ornstein, 1954) it appears that no consistent or concerted efforts have been made by the military to develop a cadre of flight instructors who agree both on the basic elements of flying proficiency (qualitatively and quantitatively) and who provide consistent and uniform courses of instruction. Quality control should be a bilateral process with as much concern for instructor differences, (which affect student-pilot achievement) as for student differences. Pilots, it appears, are assigned instructor duty principally on the basis of their experience and/or excellence as pilots.

The practice of assigning the more experienced men to duty as flight instructors is logically defensible. Presumably, they are able to pass this experience on to trainees and thereby enrich the training program. There are no data, however, to indicate that more experienced

pilots make better instructors than less experienced pilots. A small amount of evidence suggests rather that experienced pilots are no better instructors than relatively inexperienced pilots. Studies supporting this conclusion are reviewed below.

Proficiency of Instructors: An exploratory study for the Navy (Bowers, 1958) casts some doubt on the often-voiced opinion that experience in flying is, in itself, a satisfactory predictor of instructor success. Bowers analyzed available record data on certain individual differences among pilots to determine the consequences of those differences on their learning to be instructors. Using time required to complete the instructor syllabus as a criterion, he found that aviators who were younger and held relatively low rank did better in the training syllabus than did their older higher-ranking counterparts. While greater time in single-engine aircraft tended to enhance a student-instructor's performance in training, greater experience in multiengine aircraft had the opposite effect. Perhaps, the most significant finding was that experience in multiengine aircraft did not transfer to the similar yet different task of flying a single-engine training aircraft. Thus, more experienced pilots apparently have greater difficulty in acquiring instructor skills than the less experienced.

From 1955 to 1959, the U.S. Navy, owing to conditions which demanded full use of operationally qualified pilots in the Fleet, assigned a number of pilots to instructor duty immediately upon completion of flight training. The proficiency of these "plowback" instructors was subsequently compared with that of "Fleet-experienced" instructors, i.e., pilots who had served at least one tour of duty with the Fleet before being returned for instructor duty. Gallagher and Lowi (1958) used record data to evaluate differences in the quality of pilots produced by 73 plowback versus 62 Fleet-experienced instructors. Data were collected on all those students who received A-Stage training from any one of the subject instructors during the first six months in which that particular instructor taught. Only those students who had a single instructor for the entire A-Stage were used (n = 222 for plowbacks and n = 165 for Fleet-experienced). Students' scores on various measures of proficiency (e.g., number of accidents, number of downs, flight check grades, final presolo grade, etc.) were compared. Although the measures used appeared to indicate superior achievements of plowback students compared to Fleet-returnee students, all could be attributed to chance except final presolo grade, which was attributed to greater grading leniency by the plowback instructors. The authors concluded that within the limits of the primary syllabus (SNJ) that either class of pilot instructor could be used effectively as operational requirements dictated.

A subsequent Navy report (Johnson & Berkshire, 1960) was concerned with whether plowback instructors "differed" in other respects from Fleet-experienced instructors. To check on suggestions that student attitudes towards relatively inexperienced plowback instructors were not as good as toward instructors who were older, had more rank, and who could teach with authority based on actual squadron experience, ratings of 68 plowback instructors made by 121 students were compared to ratings of 43 Fleet-returnee instructors made by 86 of their students. No significant differences were reflected. An additional test of attitudes was made at the end of basic training. It was assumed that the students of the experienced instructors would be better motivated to extend their obligated military service in exchange for choice of type of advanced training. In terms of percentage of men extending, however, there were no significant differences. Training accidents were also assessed. During the first four months of 1958, plowbacks (34 percent of the pilot training instructors) accounted for 53 percent of all training accidents. However, the plowback instructors had less rank than Fleet-returnees and consequently fewer administrative duties. They, therefore, flew more training hours. Analysis of accidents by actual number of hours flown showed that again there were no significant differences between the two types of instructors. However, Johnson and Berkshire did find that while the first tour instructors were satisfactory in the instructional role, this experience affected their own subsequent performance in the Fleet. Of the plowback instructors reassigned to Fleet squadrons, more than twice the normal number were judged by their squadron commanders to be unsatisfactory both as pilots and as officers. The authors associated these differences with the fact that plowbacks spent upwards of one year after designation in flying basic training-type aircraft on instructional missions under VFR conditions which may have constituted negative conditioning for operational flying. As for officer quality, duties attached to the instructional job were considered to be poor preparation for those encountered in the Fleet squadron.

The studies reviewed above support the generalization that there is no basis for the notion that experienced pilots make better instructors than relatively inexperienced pilots. There are apparently no real differences in the primary flight training success of students taught by either class of pilot. Whether proficiency differences are instilled which may manifest themselves in later operational flying is unknown. Student attitudes toward relatively inexperienced instructors do not differ from those toward more experienced instructors, and training accidents are not significantly greater for the less experienced instructor. In fact, instructor pilots as a group may be safer pilots than operational aviators. Although the data are old, a study by Zeller and

Burke (1956), evaluated the proficiency of USAF instructor pilots in trainer-type aircraft by comparing their aircraft accident record with that of all USAF pilots during the period from 1 January 1952 through 31 December 1954. They found that instructor pilots consistently had an accident rate lower than the USAF average in comparable type aircraft, and that even when held responsible for the accidents of their students, their accident rate remained low. While all available evidence supports the generalization that it is not necessary to have a high level of experience as a pilot in order to be a successful instructor, it does appear that the careers of pilots assigned instructor duty prior to operational assignments may be adversely affected by this experience. Unfortunately no data were available regarding the subsequent performance of the Fleet-returnee instructors upon their reassignment to the Fleet to determine how their experience might have affected them.

The preceeding studies indicate that a high level of pilot experience is not necessary for a pilot to be effective as an instructor. How much the experienced pilot may enrich the training program by his experience or contribute to the proficiency of his students for later operational flying is, however, unknown and data bearing on this point should be collected.

The contribution that the flight instructor does make to training is not completely understood. We have little knowledge of what he does as a teacher and how important he is to the training process. Studies should be addressed toward precise definition of his role and determination of his function (which may at times include serving as a safety pilot for well-motivated trainees and competent trainees).

Effort is also needed to clarify the role of the simulator instructor for effective training. How does he affect learning (and subsequent transfer) of skills? Previous research (Williams & Flexman, 1949) has suggested that the use of simulators for training is importantly affected by the flight instructor who may either capitalize on or negate the value of simulator training by his opinions about what the student should learn. Instructor variables may also interact importantly with the fidelity of simulation in training devices. The nature of this interaction appears to be such that devices of low fidelity require both an increased quality and quantity of instruction to overcome inherent limitations.

Physiological Indoctrination

Current flight missions are conducted in increasingly hostile environments and pilots are encountering events heretofore never

experienced. It is thus mandatory to acquaint pilots with these events during the course of training in order to minimize the possibility of performance disruption during certain emergency situations. What this amounts to narrows down to an indoctrination of pilots in a variety of physiological sensations and unfamiliar spatial relationships, as well as to preparing him to cope successfully with situations where his life is dependent on the learning of procedures involving escape and survival equipment.

Several techniques and devices are discussed here which provide this kind of indoctrination to the trainee, i.e., acquainting him with the onset, severity, and duration of "unusual" or infrequently experienced events in the air. These involve spatial disorientation, simulated ejection from an aircraft, and the generation of high-intensity light.

Spatial Disorientation: Spatial disorientation, the incorrect perception of attitude and position in relation to the earth, is one of the major causes of aircraft accidents. In an attempt to minimize this problem, the Tactical Air Command in 1961 instituted a spatial disorientation indoctrination program (TAC Regulation 60-13) requiring TAC crewmembers to receive spatial disorientation indoctrination once a year. This program includes lectures, showing of a training film, and actual inflight demonstrations of the phenomenon. This, and similar requirements, have been the impetus for the development of means for coping with disorientation during flight.

At the request of TAC a spatial disorientation demonstrator was developed by the U.S. Air Force School of Aerospace Medicine (Lewis, Whitmore, Harris, & McDougall, 1965). The device can be rotated about three axes simultaneously or separately, and at the same time can rotate about a track at velocities up to 15 rpm. It is capable of producing illusions such as the sensations of climbing or diving while turning, tilt, reversal of motion, the Coriolis effect, inside and outside loops, Immelman turns, skids and spins. A preliminary demonstration of the device to 50 pilots at Luke Air Force Base, Arizona, indicated complete pilot acceptance. The sample of pilots agreed that it produced realistic illusions of flying through the various simulated maneuvers which had been programmed. Instructor pilots felt that the trainer could be used to replace actual inflight demonstrations since it could demonstrate essentially the same illusions. A program for the use of the device, based on pilot experiences with vertigo and spatial disorientation, has been developed (Doppelt, 1965), but no wide-scale employment of the technique in pilot training appears to have occurred. Although the device was developed for demonstrating spatial disorientation, it also appears

useful for providing training in motion cues, i.e., as an adjunct to flight simulators. Presumably, it may also serve as a pilot-selection device, in a manner similar to that of the "Whirlymite" Trainer (see pp. 155-156).

Current research is also being directed towards the possibility of training the vestibule of the inner ear as a means of reducing the incidence of both spatial disorientation and motion sickness in pilots. Gillingham (1965) presents a review and analysis of the literature which supports the position that the vestibular system is capable of being trained. Lewis et al. (1965) believe that the spatial disorientation demonstrator could be used as a means of developing resistance to spatial disorienting stimuli either by suppression of the vestibular response from repeated exposures or by overlearning. Owing to the operational importance to the Air Force of factors such as vertigo, spatial disorientation, motion sickness, and so forth, continuing research in this area is a necessity.

<u>Experiencing Ejection</u>: The value of ejection seat tower rides, was dramatically illustrated several years ago by an analysis of accident statistics for the period 1 January 1957 to 30 June 1959. This analysis cited by Beer et al., 1961, p. 30) related accident data to the type of ejection training given, and found that of the 716 ejections during this period, 123 resulted in fatalities. Of the 716 aircrewmembers who ejected, 295 (41 percent) had had at least one tower ride prior to ejection. In this group of 295, there were no fatalities, and 207, or 70 percent of this group escaped without injury of any kind. Despite the fact that pilots complain about this somewhat severe experience, tower rides which acquaint the pilot with the physiological sensations and subjective experiences associated with ejection are of obvious worth.

<u>Flash Blindness Training</u>: An indoctrination device for demonstrating sudden and complete vision loss and the concomitant performance decrement that occurs when pilots are exposed to high-intensity light such as generated by nuclear bursts has recently been developed (see p. 110). As yet, there has been no wide-scale employment of this type of training device. In spite of the importance of this kind of experience in current mission requirements, flash blindness indoctrination has not been given the attention needed.

In summary, wherever physiological factors affect the pilot's ability to maintain an adequate sustained level of performance, increased research effort should be directed. Where indicated, techniques and devices for providing exposure to, and experience with, unusual events

during flight should be developed, to be used compatibly with existing training programs.

Escape Training

Training for escape from disabled aircraft represents a growing concern to the Air Force. As aircraft fly in an environment increasingly hostile to human survival, aircrew must be proficient to the point of near-automaticity in the initiation and execution of ejection procedures. Pilots, however, are often left to their own devices for learning and mastering ejection procedures for specific aircraft, in the sense of being required to read and retain information contained in Flight Manuals, Technical Orders, manufacturers' instructions, and the like. Studies and surveys have indicated that training for escape has been inadequate both in content and in frequency.

An early study (DeGaugh & Keller, 1957) demonstrated that only 4 percent of a randomly selected sample (n = 164) of SAC aircrew members at four Air Force bases could complete ejection procedures (performance test) without error. Yet of the same group, about 55 percent performed a satisfactory preflight and demonstrated satisfactory knowledge of the operation of the system.

A survey of escape training in the Air Force (Beer et al., 1961) noted that the DeGaugh and Keller report had little discernible effect on ejection training procedures three years after the study. This survey indicated that there was a lack of standardized regular escape training programs in the Air Force and that the training media used were deficient in both quality and quantity. They concluded that, in general, personnel know when to eject but hesitated to take action because of an inadequate knowledge of procedures and an anxiety produced by unfamiliarity with the ejection experience (see also pp. 203-205). Beer et al. recommended specific ways for improving escape training with an emphasis on "confidence" training for reducing fear and anxiety about ejection.

Danaher and Sylvestro (1961) surveyed escape training for the Navy at approximately the same time as Beer et al. did for the Air Force. Many similar conclusions were reached, i.e., escape training was inadequate. This study, however, emphasized the need for greater training in the preparation phase of the escape situation, i.e., detection, diagnosis, decision, and remedial action.

Both studies concluded that pilots need more adequate and more frequently recurring ejection training than is customarily given. Three kinds of training are needed: confidence training, procedural training, and decision-making training. Indoctrination training such as provided by tower rides (see p. 205) can help to increase the trainee's confidence in the ejection system by giving him information and experience about an unknown and untried event, thereby reducing the fear and anxiety associated with ejection. Procedural training is a necessity because of the complex procedures involved and the nonstandardization of these procedures for the various ejection seats in use today. A high degree of overlearning in this operation is mandatory to insure correct response (in sequence and timing) in this extreme-stress situation. The increasing employment of jettisonable cockpits (e.g., F-111, XB-70), and increasing simplicity in escape actuating systems, decision-making training regarding when to eject will increase in importance. Such training should emphasize instruction in those emergencies which, in fact, do require escape from the aircraft and should also emphasize the safe performance envelopes.

The issue for research is to determine the current effectiveness of training as a prelude to improving its methods and mechanics where warranted. Actual performance testing rather than simple assessment of knowledge about procedures seems clearly indicated.

Sensory Training

It is well known that sensory modalities other than vision and hearing provide relatively uncluttered information channels for imparting information to individuals and for supplementing or perhaps, in some cases, supplanting the more used receptor channels. The most important of these is tactile stimulation and the use of the skin for training purposes. Not much attention has been given to this sensory channel for training in the job of flying. It is mentioned here briefly to call attention to its training possibilities.

Investigations have shown that the skin has considerable sensory ability (Mowbray & Gebhard, 1958), and the possibility of using this ability for communication purposes has been systematically explored at the University of Virginia. Geldard (1962) has summarized fourteen years of research of the Virginia Cutaneous Project investigating the vibratory sensitivity of the skin as a means of communication. The majority of these studies have been concerned with understanding of the basic nature of cutaneous vibratory phenomena. Some research has demonstrated that subjects could be trained to receive messages via the

skin at a reception rate exceeding the limits for Morse code. Howell[29] used a set of open code positions that could be keyed to letters of the alphabet. The set was composed of five separate loci, three intensities and three durations of vibration. To these 45 positions, Howell attached letters of the alphabet and common short words (such as "the" and "and"). Three subjects were trained to identify the letters corresponding to the selected code positions by a modification of a method used to teach Morse code in World War II. The subject was presented a warning signal and then the pattern. In early stages of training, the subject was told the correct letter and immediately given the same pattern to assure reinforcement. When subjects reached an acceptable level of performance on letters, they were given words of three to eight letters in length. By means of an improved sending device and further practice, the rate of transmission could be increased to a peak performance for one subject of 38 words per minute. The subject could maintain this level with no appreciable deterioration in performance. Continued explorations for training in this area are encouraged.

Tactile stimulation as a training medium, has similarly received little attention in the literature. Engelbart and Sorenson (1965) taught subjects to operate a five-key chord keyset for transmitting alphanumeric characters, and the effects of visual prompts versus tactile prompts (air jets stimulating the fingers) were compared for training efficiency. Mean differences in terminal response accuracy and response speed were not significant between groups.

The value of tactile stimulation as a means of training operator control and psychomotor tasks appears promising, but it is a relatively unexplored area and continued research is needed to determine its value both as a primary and a secondary training medium. The specific value to pilot training afforded by the modality of touch is worth considering.

Attitudes Toward Equipment

How well an operational pilot performs his total job depends upon his attitude toward performing it as much as upon anything else. His willingness and readiness to respond in demanding situations or to adjust

---

[29] Howell, W.C., 1956. Training on a Vibratory Communications System, Unpublished Master's Thesis, University of Virginia, Charlottesville; cited by Geldard, (1962).

to the requirements of the moment are important determiners of performance, whether that performance refers to critical inflight segments or to relationships with indigenous forces of a friendly country. The need to explore procedures for developing and maintaining positive attitudes and professionalism in operational pilots has been discussed earlier.

One line of inquiry has investigated pilot performance as a function of attitude toward equipment. This is an important issue for operations as well as for training. For example, low-altitude, high-speed flight requires that the pilot have confidence in his displays for terrain following and terrain avoidance and in his inertial navigation system if he is to accomplish this portion of flight successfully. Similarly, motivational similarity in the simulator is a requirement for training. This refers to the similarity in the trainee's attitude in the simulator to the feeling experienced in the actual aircraft. This can be only partially achieved, and because of its intangible nature, it is often ignored, although it forms a critical aspect of the training simulator. Unless the trainee captures some of the feeling associated with the events simulated, much of the value of even the finest operational simulator may be lost.

Two studies have been concerned with manipulating attitudes toward equipment in order to enhance the acceptance of the item(s). An unpublished University of Illinois study[30] investigated the effects of subjects' attitudes toward Link Trainer upon subsequent transfer to the T-6 aircraft. One experimental group was instilled with high confidence in the trainer by the instructor's continual stressing of the proved value of the device and the fact that previous research had shown the Link to be important in improving flying skills. The other experimental group acquired low confidence through instructor depreciation of the device. Both experimental groups learned to fly the same maneuvers in the T-6 aircraft that they had previously learned in the trainer. Trainer groups' performance was compared with a control group which had no trainer practice. The number of subjects used in each group was small (n = 4) and reliable conclusions were not reached. Analysis of the data trends, indicated that transfer was a function of total practice time in the trainer and that the attempt to instill confidence in a trainer either by instructional or apparatus techniques is a questionable procedure.

---

[30] Solarz, A.K., Matheny, W.G., Dougherty, Dora J., and Hasler, S.G. The Effect of Attitude Toward Link Training Upon Performance in the Aircraft, 1953. Cited in Adams, (1957).

Manipulation of attitudes may, however, be of considerable importance in persuading pilots to accept and use new equipment. Matheny and Berger (1964) conducted an exploratory study to investigate techniques for encouraging changes in pilots' attitudes toward new equipment. The effect of information handouts versus actual performance with new equipment as a means of changing preferences for an altimeter was investigated on a sample of fifty military helicopter pilots. The findings suggested that successful performance with the new equipment led to greater and more consistent changes in preference than did simply being given information about the new equipment.

Research investigating techniques for changing or enhancing pilots' attitudes towards available or new equipment should be extended. System effectiveness is obviously dependent upon pilot acceptance, confidence in, and use of new equipment. The simple training expedient of providing successful experiences with new equipment may be highly effective for enhancing operational capability.

REFERENCES

Abma, J. S. Theory and research in programmed instruction. AMRL Memorandum P-74, 6570th Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1964.

Adams, J. A. Some considerations in the design and use of dynamic flight simulators. AFPTRC-TN 57-51, Air Force Personnel and Training Research Center, Randolph AFB, Texas, 1957.

Adams, J. A. Part trainers. In G. Finch (Ed.) Educational and training media: a symposium. Publication 789, National Research Council, National Academy of Sciences, Washington, D.C., 1960.

Adams, J. A. Human tracking behavior. Psychol. Bull., 1961, 58, 55-79.

Adams, J. A. Motor skills. In P. R. Farnsworth (Ed.) Annual review of psychology. Palo Alto, Calif.: Annual Reviews, Inc., 1964. Pp. 181-197.

Adams, J. A., & Hufford, L. E. Effects of programmed perceptual training on the learning of contact landing skills. NAVTRADEVCEN 297-3, U.S. Naval Training Device Center, Port Washington, N.Y., 1961.

Adams, J. A., Hufford, L. E., & Dunlop, J. M. Part- versus whole-task learning of a flight maneuver. NAVTRADEVCEN 297-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1960.

Adams, J. A., & McAbee, W. H. A program for a functional evaluation of the GAM-83 Melpar Trainer. APGC-TN 61-41, Air Proving Ground Center, Eglin AFB, Florida, 1961.

Adams, J. A., & Webber, C. E. A Monte Carlo method of tracking behavior. Human Factors, 1963, 5, 81-102.

Adams, J. J., & Bergeron, H. P. Measured variations in the transfer function of a human pilot. NASA RP-206, National Aeronautics and Space Administration, Washington, D.C., 1964.

Adams, J. J., Kincaid, J. K, & Bergeron, H. P. Determination of critical tracking tasks for a human pilot. NASA Technical Note D 3242, National Aeronautics and Space Administration, Washington, D.C., Feb. 1966.

Adams, O. S., & Chiles, W. D. Human performance as a function of the work-rest cycle. WADD-TR 60-248, Wright Air Development Division, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1960.

Aerospace Medical Research Laboratories. Considerations in the design of automatic proficiency measurement equipment in simulators. AMRL Memorandum P-40, Wright-Patterson AFB, Ohio, June 1963.

Alexander, L. T., & Cooperband, A. S. System training and research in team behavior. Technical Memorandum 2581, System Development Corporation, Santa Monica, Calif., Aug. 1965.

Alluisi, E. A., Chiles, W. D., & Hall, T. J. Combined effects of sleep loss and demanding work-rest schedules on crew performance. AMRL-TDR 64-63, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1964.

Alluisi, E. A., Hall, T. J., Hawkes, G. R., & Chiles, W. D. Human performance during confinement: final report. Lockheed Georgia Company, U.S. Air Force contract AF33(616)-7604-M-4, Nov. 1962.

Ambler, Rosalie K., & Berkshire, J. R. The effects of four instructional flights during pre-flight indoctrination. Special Report No. 60-5, U.S. Naval School of Aviation Medicine, U.S. Naval Aviation Medical Center, Pensacola, Florida, 1960.

Ammons, R. B., et al. Long-term retention of perceptual-motor skills. J. exp. Psychol., 1958, 55, 318-328.

Annett, J. The role of knowledge of results in learning: a survey. NAVTRADEVCEN 342-3, U.S. Naval Training Device Center, Port Washington, N.Y., 1961.

Bartlett, F. Fatigue in the air pilot. FPRC 488, Flying Personnel Research Committee, Institute of Aviation Medicine, Royal Air Force, Farnborough, England, 1942.

Bartley, S. H., & Chute, E. Fatigue and impairment in man. New York: McGraw-Hill, 1947.

Beer, M., Jayson, R. M., Carter, V. E., & Kresse, F. H. Survey of escape training in the Air Force. WADD-TR 60-792, Wright Air Development Division, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1961.

Benenati, A. T., Hull, R., Korobow, N., & Nienaltowski, W. Development of an automatic monitoring system for flight simulators. AMRL-TDR 62-47, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, May 1962.

Bergum, B. O., & Klein, C. I. A survey and analysis of vigilance research. Research Report 8, Human Resources Research Office, Alexandria, Va., 1961.

Berkshire, J. R. The Pensacola student prediction system: progress report. Special Report 65-1, U.S. Naval School of Aviation Medicine, Pensacola, Florida, Jan. 1965.

Berkshire, J. R., & Ambler, Rosalie K. The value of indoctrination flights in the screening and training of naval aviators. Aerospace Med., 1963, 34 (5), 420-423.

Berkun, M. M. Performance decrement under psychological stress. Human Factors, Feb. 1964, 21-30.

Berkun, M. M., Bialek, H. M., Kern, R. P., & Yagi, K. Experimental studies of psychological stress on man. Psychol. Monogr., 1962, 76, No. 15 (Whole No. 534).

Besco, R. V. The effects of cockpit vertical accelerations on a simple piloted tracking task. North American Aviation Report No. NA 61-47, Los Angeles, Calif., April 1961.

Billings, C. E., Eggspuehler, J. J., & Gerke, R. Studies of pilot performance in the flight environment. Report No. RF-1857-2, Ohio State Research Foundation, contract DA 49-193-MD-2615, U.S. Army Research and Development Command, Washington, D.C., May 1966.

BioTechnology, Inc. Instructor's guide for flash blindness indoctrination trainer, device 18F22. NAVSO-P 2953, U.S. Naval Training Device Center, Port Washington, N.Y. 1966.

Birmingham, H. P., & Taylor, F. V. A design philosophy for man-machine control systems. Proc. I.R.E., 1954, 42, 1748-1758.

Boguslaw, R., & Porter, E. H. Team functions and training. In R. Gagne (Ed.) Psychological principles in system development. New York: Holt, Rinehart, & Winston, 1963. Pp. 387-416.

Bowen, H. M., Bishop, E. W., Promisel, D., & Robins, J. E. Study, assessment of pilot proficiency. NAVTRADEVCEN 1614-1, U.S. Naval Training Device Center, Port Washington, N.Y., Aug. 1966.

Bowen, H. M., Hale, A., & Kelley, C. R. Tracking training V: field study of the training effectiveness of the general vehicular research tool. NAVTRADEVCEN-TR 955-1, U.S. Naval Training Device Center, Port Washington, N.Y., Dec. 1962.

Bowers, N. D. Factors related to achievement in an instructor's training syllabus. Special Report No. 58-4, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1958.

Boyle, D. J., & Hagin, W. V. The light plane as a pre-primary selection and training device: I. Analysis of operational data. HRRC-TR 53-33, Human Resources Research Center, Air Research and Development Command, Goodfellow AFB, Texas, 1953.

Briggs, G. E., Fitts, P. M., & Bahrick, H. P. Learning and performance in a complex tracking task as a function of visual noise. Research Report AFPTRC-TN 56-67, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1956.

Briggs, G. E., & Naylor, J. C. The relative efficiency of several training methods as a function of transfer task complexity. J. exp. Psychol., 1962, 64, 505-512.

Briggs, G. E., & Naylor, J. C. Team versus individual training, training task fidelity, and task organization effects on transfer performance by three-man teams. J. appl. Psychol., 1965, 49 (6), 387-392.

Briggs, G. E., Naylor, J. C., & Fuchs, A. H. Whole versus part training as a function of task dimensions. NAVTRADEVCEN-TR 950-2, U.S. Naval Training Device Center, Port Washington, N.Y., 1962.

Briggs, G. E., & Wiener, E. L. Fidelity of simulation: I. Time sharing requirements and control loading as factors in transfer of training. NAVTRADEVCEN-TR 508-4, U.S. Naval Training Device Center, Port Washington, N.Y., 1959.

Briggs, G. E., & Wiener, E. L. Influence of time sharing and control loading on transfer of training. J. appl. Psychol., 1966, 50, 201-203.

Broadbent, D.C. Perception and communication. New York: Pergamon Press, 1958.

Buckhout, R., Naylor, J. C., & Briggs, G. E. Effects of modified task feedback during training on performance of a simulated attitude control task after thirty days. AMRL-TDR 63-125, Behavioral Sciences Laboratory, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, Dec. 1963.

Buckhout, R., Sherman, H., Goldsmith, C. T., & Vitale, D. A. The effect of variations in motion fidelity during training on simulated low altitude flight. AMRL-TDR 63-108, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, Dec. 1963.

Buckner, D. N., & McGrath, J. J. (Eds.) Vigilance: A symposium. New York: McGraw-Hill, 1963. P. 269.

Buddenhagen, T. F., Johnson, A. B., Stephan, S. C., & Wolpin, M. P. Development of visual simulation techniques for astronautical flight training: I. High resolution television; electronic planetarium. AMRL-TDR 63-54, 6570th Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1963.

Buddenhagen, T. F., & Wolpin, M. P. A study of visual simulation techniques for astronautical flight training. WADD-TR 60-756, Wright Air Development Division, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1961.

Caro, P. W., Jr., & Isley, R. N. Changes in flight trainee performance following synthetic helicopter flight training. Professional Paper 1-66, Human Resources Research Office, Alexandria, Va., 1966.

Caroll, J. B. Research on teaching foreign languages. In N. L. Gage (Ed.) Handbook of research on teaching. Chicago: Rand McNally, 1963. P. 1218.

Carstater, E. D. Comments reported in P. H. DuBois & W. H. Manning (Eds.) Methods of research in technical training. Report of conference held at U.S. Naval Air Station, Memphis, Tenn., Oct. 1956.

Chapman, R. L., Kennedy, J. L., Newell, A., & Biel, W. C. The Systems Research Laboratory's air defense experiments. Mgmt. Sci., 1959, 5, 250-263.

Chenzoff, A. P. A review of the literature on task analysis methods. NAVTRADEVCEN-TR 1218-3, U.S. Naval Training Device Center, Port Washington, N.Y., 1964.

Chernikoff, R., Duey, J. W., & Taylor, F. V. Two dimensional tracking with identical and different control dynamics in each coordinate. U.S. Naval Research Laboratory Report 5424, U.S. Naval Research Laboratory, Washington, D.C., Nov. 1959.

Chiles, W. D. Psychological stress as a theoretical concept. WADC-TR 57-457, Wright Air Development Center, Wright-Patterson AFB, Ohio, 1957.

Coermann, R. R., Magid, E. B., & Lange, K. O. Human performance under vibrational stress. Human Factors, 1962, 4 (5), 315-324.

Cole, G., et al. Study of pilot-controller integration for emergency conditions. RTD-TDR 63-4092, Research and Technology Division, Air Force Systems Command, Wright-Patterson AFB, Ohio, 1963.

Coulson, J. E. (Ed.) Programmed learning and computer-based instruction. New York: John Wiley and Sons, Inc., 1962.

Cox, J. A., & Mullins, C. J. Evaluation of light plane training among AFROTC student officers. WADC-TN 59-43, Wright Air Development Center, Personnel Laboratory, Lackland AFB, Texas, 1959.

Cox, J. A., Wood, R. O., Jr., Baren, L. M., & Thorne, H. W. Functional and appearance fidelity of training devices for fixed-procedures tasks. HumRRO-TR 65-4, Human Resources Research Office, Alexandria, Va., 1965.

Crawford, M. P. Concepts of training. In R. Gagne (Ed.) Psychological principles in system development. New York: Holt, Rinehart, & Winston, 1962. Pp. 301-341.

Crawford, M. P., & Daily, J. T. An analysis of elementary pilot performance from instructors' comments. Amer. Psychol., 1946, 1, 292.

Creelman, J. A. Evaluation of approach training procedures. Report No. 2, Project No. NM 001-109-107, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1955. (a)

Creelman, J. A. Evaluation of a proposed instrument sequence. Part I: Acrobatic stage criteria. Special Report 55-5, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1955. (b)

Creelman, J. A. Evaluation of a proposed instrument sequence. Part II: Basic non-instrument proficiency. Special Report 55-16, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1955. (c)

Creelman, J. A. Evaluation of a proposed instrument sequence. Part III: Basic instrument and night primary proficiency. Special Report 55-18, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1955. (d)

Creelman, J. A. Evaluation of a proposed instrument sequence. Part IV: Advanced instrument proficiency. Special Report No. 56-12, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1956.

Danaher, J. W., & Sylvestro, A. Aviator escape systems training, NAVTRADEVCEN-TR 880-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1961.

Danneskiold, R. D. Objective scoring procedure for operational flight trainer performance. SDC 999-2-4, U.S. Naval Special Devices Center, Port Washington, N.Y., Feb. 1955.

Danneskiold, R. D., & Johnson, W. An evaluation of an experimental flight grading method for use in the U.S. Naval Air Basic Training Command. Contract Nonr-1162(00), Psychological Corporation, and the U.S. Naval School of Aviation Medicine, Pensacola, Florida, Aug. 1954.

Davis, D. R. Pilot error: some laboratory experiments. London, England: His Majesty's Stationery Office, 1948.

Dawkins, P. B. Programmed instruction and low altitude aerial observation. HumRRO Research Report 14, Human Resources Research Office, Alexandria, Va., 1964.

Dawkins, P. B. Automated education in the training of low altitude aerial observers. In Department of the Army Tenth Annual Research and Development Conference. Washington, D.C.: Chief of Research and Development, 1965.

Deese, J. Skilled performance and conditions of stress. In R. Glaser (Ed.) Training research and education. Pittsburgh: University of Pittsburgh Press, 1962. Pp. 199-222.

DeGaugh, R. A., & Keller, J. R. Ejection seat training in the Strategic Air Command. AFPTRC Technical Memorandum OL-TM 57-16, Air Force Personnel and Training Research Center, Randolph AFB, Texas, 1957.

DeGaugh, R. A., & Knoell, D. M. Attitudes relevant to bomber crew performance in combat. AFPTRC-TR 54-18, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas, 1954.

Demaree, R. G., Norman, D. A., & Matheny, W. G. An experimental program for relating transfer of training to pilot performance and degree of simulation. NAVTRADEVCEN-TR 1388-1, U.S. Naval Training Device Center, Life Sciences, Inc., Ft. Worth, Texas, June 1965.

Diamantides, N. D., & Cacioppo, A. J. Human response dynamics; GEDA computer application. GER-8033, Goodyear Aircraft Corporation, Akron, Ohio, Jan. 1957.

Dobbins, D. A., Tiedemann, J. G., & Skordahl, D. M. Field study of vigilance under highway driving conditions. TRN No. 118, U.S. Army Personnel Research Office, Washington, D.C., 1961.

Dobbins, D. A., Tiedemann, J. G., & Skordahl, D. M. Vigilance under highway driving conditions. Percept. mot. Skills, 1963, 16, 38.

Doppelt, F. F. A program for the spatial disorientation demonstrator. SAM-TR 65-66, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1965.

Dougherty, Dora, Houston, R. C., & Nicklas, D. R. Transfer of training in flight procedures from selected ground training devices to the aircraft. NAVTRADEVCEN-TR 71-16-16, U.S. Naval Training Device Center, Port Washington, N.Y., 1957.

Duffy, J. O., & Colgan, C. M. A system of flight training quality control and its application to helicopter training. Consulting Report, U.S. Army Aviation Research Unit, Ft. Rucker, Ala., June 1963.

Dunlap and Associates, Inc. A method for deriving job standards from system effectiveness criteria: I. Method development. Contract Nonr 4314(00), Santa Monica, Calif., 1964.

Dunlap and Associates, Inc. Development of a technique for establishing personnel performance standards, Phase II, interim report. Contract Nonr 4314(00), Santa Monica, Calif., 1965.

Eachus, H. T. Self-confrontation for complex skill training: review and analysis. AMRL-TR 65-118, 6570th Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1965.

Eakin, G. J., & Campbell, G. F. Human operator transfer function investigation. Contract No. AF33(038)-9200, Report No. TE-665-F-5, Cornell Aeronautical Laboratory, Inc., Buffalo, N.Y., Oct. 1957.

Eckstrand, G. Current status of the technology of training. AMRL-TR 64-86, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, Sep. 1964.

Edgerton, H., & Walker, R. Y. History and development of the Ohio State Flight Inventory. Part I: Early versions and basic research. Report No. 47, CAA Division of Research, CAA, Washington, D.C., 1945.

Egerman, K., Klaus, D. J., & Glaser, R. Increasing team proficiency through training. III. Decremental effects of reinforcement in teams with redundant members. Contract Nonr 2551(00), American Institute for Research, Pittsburgh, Pa., 1962.

Eliasson, A. H. The sequence of flight instruction. Army Aviation, 1961, 10 (6), 341-344.

Elkind, J. I., & Green, D. M. Measurement of time-varying and nonlinear dynamic characteristics of human pilots. ASD Technical Report 61-225, Aeronautical Systems Division, Wright-Patterson AFB, Ohio, Dec. 1961.

Elliott, E. Perception and alertness. Ergonomics, 1960, 3, 357-364.

Engelbart, D. C., & Sorensen, P. H. Explorations in the automation of sensorimotor skill training. NAVTRADEVCEN-TR 1517-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1965.

Ericksen, S. C. A review of the literature on methods of measuring pilot proficiency. Research Bulletin 52-25, Human Resources Research Center, Lackland AFB, Texas, 1952. (a)

Ericksen, S. C. Development of a light plane proficiency check to predict military flying success. Technical Report 52-6, Human Resources Research Center, Lackland AFB, Texas, Nov. 1952. (b)

Feddersen, W. E. Simulator research: validation and motion studies. In Seventh Annual Army Human Factors Engineering Conference, Oct. 1961.

Feddersen, W. E. The role of motion information and its contribution to simulation validity. Tech. Data Report No. D 228-429-001, Bell Helicopter Co., Ft. Worth, Texas, 1962.

Fiedler, F. The influence of leader-key man relations on combat crew effectiveness. Contract N S ori 07135, Technical Report No. 9, Office of Naval Research, Washington, D.C., 1958.

Finan, J. L., Finan, S. C., & Hartson, L. D. A review of representative tests used for the quantitative measurements of behavior-decrement under conditions related to aircraft flight. USAF-TR 5830, Aero Medical Laboratory, Wright-Patterson Air Force Base, Ohio, 1949.

Fitts, P. M. Factors in complex skill training. In R. Glaser (Ed.) Training research and education. Pittsburgh: University of Pittsburgh Press, 1962. Pp. 177-197.

Fitzpatrick, R. The training effectiveness of different arrangements of aircraft and simulator training: II. C-124 pilots. CRL-TM 55-13, Crew Research Laboratory, Air Force Personnel and Training Research Center, Randolph AFB, Texas, 1955.

Fleishman, E. A. The description and prediction of perceptual-motor skill learning. In R. Glaser (Ed.) Training research and education. Pittsburgh: University of Pittsburgh Press, 1962. Pp. 137-175.

Fleishman, E. A., & Hempel, W. E. The relation between abilities and improvement with practice in a visual discrimination reaction task. J. exp. Psychol., 1955, 49, 301-312.

Fleishman, E. A., & Ornstein, G. N. An analysis of pilot flying performance in terms of component abilities. J. appl. Psychol., 1960, 44, 146-155.

Fleishman, E. A., & Parker, J. F., Jr. Factors in the retention and relearning of perceptual-motor skill. J. exp. Psychol., 1962, 64, 215-226.

Flexman, R. E., Matheny, W. G., & Brown, E. L. Evaluation of the School Link and special methods of instruction in a ten-hour private pilot flight-training program. Aeronautics Bull., No. 8. Urbana, Ill.: University of Illinois Institute of Aviation, 1950.

Flexman, R., Townsend, J. C., & Ornstein, G. N. Evaluation of a contact flight simulator when used in an Air Force primary pilot training program: Part I: Over-all effectiveness. AFPTRC-TR 54-38, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1954.

Forgays, D. G., & Irwin, I. A. Measures of combat crew performance used in B-29 training. Technical Report 52-14, Human Resources Research Center, Air Research and Development Command Randolph AFB, Texas, 1952.

Franklin, Margaret E., & Whittenburg, J. A. Research on visual target detection: Part I. Development of an air-to-ground detection/identification model. HSR-RR 65/4-Dt, prepared by Human Sciences Research, McLean, Va., for DA Project No. 1R12030D035, Contract DA-31-124-ARO-D-287, U.S. Army Human Engineering Laboratories, Aberdeen Md., 1965.

Fraser, T. M. Philosophy of simulation in a man-machine space mission system. NASA SP-102, National Aeronautics and Space Administration, Washington, D.C., 1966.

Fraser, T. M., Hoover, G. N., & Ashe, W. F. Tracking performance during low frequency vibration. Aerosp. Med., 1961, 32(9), 829-835.

Frost, G. G. A comparison between tracking with "optimum" dynamics and tracking with a simple velocity control. Technical Documentary Report No. AMRL-TDR 62-150, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, Dec. 1962.

Gabriel R. F., Burrows, A. A., & Abbott, P. E. Using a generalized contact flight simulator to improve visual time sharing. NAVTRADEVCEN-TR 1428-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1965.

Gage, N. L. (Ed.) Handbook of research on teaching. Chicago: Rand-McNally, 1963. P. 1218.

Gagne, R. M. Work decrement in the learning and retention of motor skills. In W. Floyd & A. Welford (Eds.) Symposium on fatigue. London: H. K. Lewis and Co., 1953. Pp. 155-163.

Gagne, R. M. Training devices and simulators: some research issues. Amer. Psychologist, 1954, 9, 95-107.

Gagne, R. M. Simulators. In R. Glaser (Ed.) Training research and education. Pittsburgh: Univ. of Pittsburgh Press, 1962, Pp. 223-246.

Gagne, R. M. Human functions in systems. In R. Gagne (Ed.) Psychological principles in system development. New York: Holt, Rinehart, & Winston, 1963. Pp. 35-74.

Gallagher, T. J., & Lowi, B. H. The proficiency of first tour pilots as instructors. Special Report No. 58-5, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1958.

Gardner, J., Lacey, R., Seeger, C., & Wade, J. In-flight comparison of pilot performance on a standard USAF and an experimental instrument panel. WADC Technical Report 57-270, Wright Air Development Center, Wright-Patterson AFB, Ohio, 1957.

Garfinkle, D. R., Smith, R. L., Lyman, J., & Groth, H. Performance studies on the NOTS-UCLA tracking simulator: Effects of error magnification, field of view, proprioceptive cues and control dynamics. UCLA Engineering Department, Biotechnology Lab, Techinical Report No. 63-41, April 1963.

Geldard, F. A. Virginia cutaneous project, 1948-1962. Final Report to the Office of Naval Research on Project NR-140-598. Psychological Laboratory, University of Virginia, Charlottesville, Va., 1962.

Gerhardt, L., & Johnson, A. B. Development of visual simulation techniques for astronautical flight training: II. Image assembly techniques for visual simulation of orbiting flight. AMRL-TDR 63-54, Vol. 2, 6570th Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1963.

Gibson, J. J. The perception of the visual world. Cambridge: Riverside Press, 1950.

Gibson, J. J. The optical expansion pattern on aerial locomotion. Amer. J. Psychol., 1955, 68, 480-484.

Gillingham, K. K. Training the vestibule for aerospace operations, central control of vestibular function. Aeromedical Reviews, Review 8-65, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1965.

Glanzer, M., & Glaser, R. Techniques for the study of group structure and behavior: II. Empirical studies of the effects of structure on small groups. Psychol. Bull., 1961, 58, 1-27.

Glaser, R. (Ed.) Teaching machines and programmed learning: data and directions. Washington, D.C.: National Education Association, 1964.

Glaser, R., & Klaus, D. J. Proficiency measurement: assessing human performance. In R. Gagne (Ed.) Psychological principles in systems development. New York: Holt, Rinehart, & Winston, 1963. Pp. 419-476.

Glaser, K., Klaus, D. J., & Egerman, K. Increasing team proficiency through training: II. The acquisition and extinction of a team response. Contract Nonr-2551(00), American Institute for Research, Pittsburgh, Pa., 1962.

Goldman, D. F., & Von Gierke, H. E. The effects of shock and vibration on man. No. 60-3, Lecture and Review Series, Naval Medical Research Institute, Bethesda, Md., 1960.

Goodwin, W. R. The System Development Corporation and system training. Amer. Psychologist, 1957, 12, 524-527.

Goodyear Aircraft Corporation. Human dynamics study (final report). GER Report 4750, Akron, Ohio, April 1952.

Gordon, T. A. The airline pilot: a survey of the critical requirements of his job and of pilot evaluation and selection procedures. Report No. 73, CAA Division of Research, CAA, Washington, D.C., 1947.

Gordon, T. A. The development of a standard flight-check for the airline transport rating based on the critical requirements of the airline pilot's job. Report No. 85, CAA Division of Research, CAA, Washington, D.C., 1949.

Greenwood, M., & Weybrew, B. B. Behavioral periodicity: I. Bibliography of literature pertaining to human beings. Report No. 421, U.S. Naval Medical Research Laboratory, New London, Conn., 1964.

Greer, G. D., Smith, W. D., & Hatfield, J. L. Manual of instruction on use of pilot performance description records in flight training quality control. Human Resources Research Office, Ft. Rucker, Ala., 1959.

Greer, G. D., Smith, W. D., & Hatfield, J. L. Improving flight proficiency evaluation in Army helicopter pilot training. HumRRO Technical Report No. 77, U.S. Army Aviation Human Research Unit, Ft. Rucker, Ala., May 1962.

Grumman Aircraft Engineering Corporation. F-111B task analysis: I. Crew activities. Report FZM-12-9181, Bethpage, N.Y., March 1965.

Haines, D. B. Training for culture contact and interaction skills. AMRL-TR 64-109, 6570th Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1964.

Haines, D. B., & Eachus, H. T. A preliminary study of acquiring cross-cultural interaction skills through self-confrontation. AMRL-TR 65-137, 6570th Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1965.

Hall, R. L. Group performance under feedback that confounds responses of group members. Sociometry, 1957, 20, 297-305.

Hartman, B. O. Fatigue effects in 24-hour simulated transport flight: changes in pilot proficiency. SAM-TR 65-16, USAF School of Aerospace Medicine, Brooks Air Force Base, Texas, 1965.

Haythorn, W. W. A review of research on group assembly. AFPTRC-TN 57-62, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas, 1957.

Herbert, M. J. Analysis of a complex skill. Vehicle driving. Report No. 581, U.S. Army Medical Research Laboratory, Ft. Knox, Ky., 1963.

Herbert, M. J., & Jaynes, W. E. Performance decrement in vehicle driving. Report No. 597, U.S. Army Medical Research Laboratory, Ft. Knox, Ky., 1963.

Hesson, J. M., & Thomas, F. H. Training materials for aerial observer instruction in basic visual skills. Supplement to Technical Report No. 80, Human Resources Research Office, Alexandria, Va., Oct. 1962.

Hickey, A. E., & Newton, J. M. Computer-assisted instruction: a survey of the literature. Report prepared by ENTELEK Inc., Newburyport, Massachusetts, for Contract Nonr-4757(00), Project Nr. 154-254 with Office of Naval Research, Wash., D.C., 1966.

Hixon, W. C., Harter, G. A., Warren, C. E., & Cowan, J. D. An electronic radar target simulator for air traffic control studies. WADC-TR 54-569, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, 1954.

Hoisman, A. J., & Daitch, A. M. Techniques for relating personnel performance to system effectiveness criteria: a critical review of the literature. Contract Nonr-4314(00), Dunlap and Associates, Inc., Santa Monica, Calif., 1964.

Hood, P. D. Early history and theory of integrated crew training. Conference on integrated aircrew training. WADD-TR 60-320, Wright Air Development Division, Wright-Patterson Air Force Base, Ohio, July 1960. Pp. 3-14.

Horrocks, J. E., Heermann, E., & Krug, R. E. Team training: III. An approach to optimum methods and procedures. NAVTRADEVCEN 198-3, U.S. Naval Training Device Center, Port Washington, N.Y., 1961.

Horrocks, J. E., Krug, R. E., & Heermann, E. Team training: II. Individual learning and team performance. NAVTRADEVCEN 198-2, U.S. Naval Training Device Center, Port Washington, N.Y., 1960.

Houston, R. C., Smith, J. F., & Flexman, R. E. Performance of student pilots flying the T-6 aircraft in primary pilot training. AFPTRC-TR 54-109, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas, 1954.

Hovland, C. I. Human learning and retention. In S. S. Stevens (Ed.) Handbook of experimental psychology. New York: Wiley, 1951. Pp. 613-689.

Hudson, E. M. Adaptive training and nonverbal behavior. NAVTRA-DEVCEN 1395-1, U.S. Naval Training Device Center, Port Washington, N.Y., July 1964.

Hufford, L. E., & Adams, J. A. The contribution of part-task training to the relearning of a flight maneuver. NAVTRADEVCEN-TR 297-2, U.S. Naval Training Device Center, Port Washington, N.Y., 1961.

Jackson, E. F. Aircrew fatigue in long range maritime reconnaissance: pilot performance. FPRC 907.2, Flying Personnel Research Committee, Institute of Aviation Medicine, Royal Air Force, Farnborough, England, 1956.

Jerison, H. J., & Pickett, R. N. Vigilance: a review and re-evaluation. Human Factors, 1963, 5 (3), 211-238.

Johnson, J. H. Effects of training plane type on jet pilot performance. Aerospace Med., 1962, 33 (9), 1103-1110.

Johnson, J. H., & Berkshire, J. R. The use of newly designated aviators as instructors. Special Report No. 60-6, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1960.

Johnston, W. A. Transfer of team skills as a function of type of training. J. appl. Psychol., 1966, 50 (2), 102-108.

Jolley, O. B. A summary of prior research on integrated contact/ instrument flight training. Staff Memorandum, Project INTACT, Human Resources Research Office, Alexandria, Virginia, 1958.

Jones, R. L., & Lindsey, J. F. Human factors aspects of low altitude flight: a sample of fighter pilot attitudes and altitude estimates. APGC-TR 65-35, Air Proving Ground Center, Eglin Air Force Base, Florida, May 1965.

Jones, R., Milton, J., & Fitts, P. M. Eye fixations of aircraft pilots: IV. Frequency, duration and sequence of fixations during routine instrument flight. USAF Technical Report 5975, Wright-Patterson AFB, Ohio, 1949.

Joska, J. S. Effect of aircraft speed on low altitude acquisition of ground targets. Technical Report 65-34, Air Proving Ground Center, Eglin AFB, Florida, May 1965.

Karlin, L. Psychological study of motor skills: Phase I. NAVTRADEVCEN-TR 558-1, U.S. Naval Training Device Center, Port Washington, N.Y., July 1960.

Katona, G. Organizing and memorizing: studies in the psychology of learning and teaching. New York: Columbia University Press, 1940.

Kelley, C. R. The predictor instrument: final report and summary of activities during 1961. Contract Nonr-2822(00), Office of Naval Research, Washington, D.C., January 1962.

Kelley, C. R. Self-adjusting vehicle simulators. Contract Nonr-2822(00), Dunlap and Associates, Darien, Conn., Dec. 1963.

Kelley, C. R., Bowen, H. M., Ely, J. H., & Andreassi, J. L. Tracking training: III. Transfer of training. NAVTRADEVCEN-TR 1908-00-3, U.S. Naval Training Device Center, Port Washington, N.Y., Jan. 1960.

Kennedy, J. L. The system approach: organizational development. <u>Human Factors</u>, Feb. 1962, 25-52.

Kibler, A. W. The relevance of vigilance research to aerospace monitoring tasks. Human Factors, 1965, <u>7</u> (2), 93-100.

Kinkade, R. G. Augmented feedback and tracking skill. NAVTRADEVCEN-TR 508-3, U.S. Naval Training Device Center, Port Washington, N.Y., Oct. 1959.

Klier, S., & Linskey, J. W. Selected abstracts from the literature on stress. NAVTRADEVCEN 565-1, U.S. Naval Training Device Center, Port Washington, N.Y., Nov. 1960.

Knoell, D. M. Relationships between attitudes of bomber crews in training and their attitudes and performance in combat. AFPTRC-TN 56-49, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1956.

Knoop, Patricia A. Programming techniques for the automatic monitoring of human performance. AMRL-TR 66-16, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, Apr. 1966.

Knowles, W. B. Operator loading tasks. Human Factors. Apr. 1963, 155-161.

Krendel, E. S., Bloom, J. W. The natural pilot model for flight proficiency evaluation. NAVTRADEVCEN-TR 323-1, U.S. Naval Training Device Center, Port Washington, N.Y., April 1963.

Krumm, R. L. Current research on integrated crew training: conference on integrated aircrew training. WADD-TR 60-320, Wright Air Development Division, Wright-Patterson AFB, Ohio, July 1960. Pp. 15-36.

Krumm, R. L., & Farina, A. J. Effectiveness of integrated flight simulator training in promoting B-52 crew coordination. AMRL-TDR 62-1, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1962.

Kurtzberg, J. M. Dynamic task scheduling in flight simulators. AMRL-TDR 63-17, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, Feb 1963.

Lamont, J. N. Annotated bibliography on flight simulators. HR-REP No. 68, Directorate of Biosciences Research, Human Resources Research Section, Ottawa, Canada, 1960.

Lanzetta, J. T., & Roby, T. B. Group learning and communication as a function of task and structure demands. J. abnorm. soc. Psychol., 1957, 55, 121-131.

Lazarus, R. S., Deese, J., & Osler, Sonia F. The effects of psychological stress upon performance. Psychol. Bull., 1952, 49, 295-317.

Lewis, R. E. F. Pilot performance during low speed, low level navigation. DRML Report 248-1, Defense Research Medical Laboratories, Toronto, Ontario, Canada, 1961.

Lewis, R. E. F., & de la Riviere, W. E. A further study of pilot performance during low speed, low level navigation. DRML Report 248-2, Defense Research Medical Laboratories, Toronto, Canada, 1962.

Lewis, S. T., Whitmore, H. B., Harris, L., & McDougall, D. A. Spatial disorientation demonstrator. SAM-TR 65-7, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1965.

Licklider, J. C. R. Quasi-linear models in the study of manual tracking. In R. D. Luce (Ed.) Developments in mathematical psychology. Glencoe, Illinois: The Free Press, 1960. Pp. 171-280.

Lincoln, R. S., & Mangelsdorf, J. E. An integrated approach to evaluating the performance capabilities and physiological state of spacecraft crews. Paper presented at Symposium on Human Physiological and Performance Determinants of Manned System Design, San Fernando Valley State College, Calif., 14 April 1965.

Lindquist, O. H., & Gross, R. L. Human engineering man-machine study of a weapon system. Report R-ED 6094, Minneapolis-Honeywell Regulator Co., Aeronautical Division, Minneapolis, Minnesota, Oct. 1958.

Longo, A. A., & Doll, R. E. An appraisal of psychological stress. Aerospace Med., 1962, 33 (9), 1111-1116.

Lumsdaine, A. A. Design of training aids and devices. In J. D. Folley, Jr. (Ed.) Human factors methods for system design. Pittsburgh: American Institute for Research, 1960.

Lumsdaine, A. A., & Glaser, R. Teaching machines and programmed learning: a source book. Washington, D.C.: National Education Assn., 1960.

Lumsdaine, A. A., & May, M. A. Mass communication and educational media. Annu. Rev. Psychol., May 1965, 475-534.

Lybrand, W. A., Havron, M. D., Gartner, W. B., Scarr, H. A., & Hackman, R. C. Simulation of extra-cockpit visual cues in contact flight transition trainers. AFPTRC-TR [illegible]-11, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas, 1958.

Mahler, W. R., & Bennett, G. K. Psychological studies of advanced naval training: evaluation of operational flight trainers. TR-SDC 999-1-1, USN, Office of Naval Research, Special Devices Center, Port Washington, N.Y., 1950.

Margulies, S., & Eigen, L. D. (Eds.) Applied programmed instruction. New York: John Wiley and Sons, 1962.

Martorana, S. V. The Air Force OAR contribution to programmed instruction. OAR 64-10, USAF, Office of Aerospace Research, Washington, D.C., 1964.

Mast, T. M., & Heimstra, N. W. Effects of fatigue on vigilance performance. J. engng. Psychol., 1964, 3 (3), 73-79.

Matheny, W. G., Williams, A. C., Jr., Dougherty, Dora, & Hasler, S.G. The effect of varying control forces in the P-1 Trainer upon transfer of training to the T-6 aircraft. HRRC-TR 53-31, Human Resources Research Center, Goodfellow AFB, Texas, 1953.

Mayo, G. D., & Longo, A. A. Training time and programmed instruction. J. appl. Psychol., 1966, 50 (1), 1-4.

McFarland, R. The psychological aspects of flying with special reference to problems of selection. Int. Clin., 1942, 1, Series 5, 17-35.

McGeoch, J. A., & Irion, A. L. The psychology of human learning. (Rev. ed.) New York: Longmans, Green and Co., 1952.

McGrath, J. J., & Borden, G. J. Geographic orientation in aircraft pilots: a problem analysis. Technical Report 751-1, Human Factors Research, Inc., Los Angeles, Calif., 1963.

McGrath, J. J., Harabedian, A., & Buckner, D. N. Review and critique of the literature on vigilance performance. TR-1, Contract Nonr-2649(00). Human Factors Problems in Anti-Submarine Warfare. Human Factors Research, Inc., Los Angeles, Calif., 1959.

McGrath, J. J., Osterhoff, W. E., & Borden, G. J. Geographic orientation in aircraft pilots: experimental studies of two cartographic variables. TR 751-3, Human Factors Research, Inc., Los Angeles, Calif., Nov. 1964.

McGrath, J. J., Osterhoff, W. E., & Borden, G. J. Geographic orientation in aircraft pilots: chart scale and pilot performance. Technical Report 751-4, Human Factors Research, Inc., Los Angeles, Calif., July 1965.

McGrath, J. J., Osterhoff, W. E., Seltzer, M. L., & Borden, G. J. Geographic orientation in aircraft pilots: methodological advancement. Technical Report 751-5, Human Factors Research, Inc., Los Angeles, Calif., Oct. 1965.

McIntosh, B. B., Milton, J. L., & Cole, E. L. Pilot performance during extended periods of instrument flight. AF-TR 6725, Wright Air Development Center, Aeromedical Laboratory, Wright-Patterson AFB, Ohio, 1952.

McRuer, D., Graham, D., Krendel, E., & Reisener, W., Jr. Human pilot dynamics in compensatory systems: theory, models, and experiments with controlled element and forcing function variations. Technical Report Nr. AFFDL-TR 65-15, Air Force Flight Dynamics Laboratory, Wright-Patterson AFB, Ohio, July 1965.

McRuer, D. T., & Krendel, E. S. Dynamic response of human operators. USAF Wright Air Development Center Report 56-524, Wright-Patterson AFB, Ohio, Oct. 1957.

Melton, A. W. Retention of tracking skills. Contract No. DA-49-007-MD-1020, U.S. Army Medical Research and Development Command, Office of the Surgeon General, Washington, D.C., Sep. 1964.

Mengelkoch, R. F., Adams, J. A., & Gainer, C. A. The forgetting of instrument flying skills as a function of the level of initial proficiency. NAVTRADEVCEN 71-16-18, U.S. Naval Training Device Center, Port Washington, N.Y., 1958.

Messick, S. Personality structure. In P. Farnsworth, O. McNemar, & Q. McNemar (Eds.) Annu. Rev. Psychol., 1961, 12, 93-128.

Miller, J. W. (Ed.) Visual, display, and control problems related to flight at low altitude. ONR Symposium Report ACR-95, Office of Naval Research, Washington, D.C., 3-5 March 1964.

Miller, N. (Ed.) Psychological research on pilot training. Army Air Force Aviation Psychology Program Research Report No. 8. Washington, D. C.: U.S. Government Printing Office, 1947.

Miller, R. B. Task and part-task trainers and training. WADD-TR 60-469, Wright Air Development Division, Behavioral Sciences Laboratory, Wright-Patterson AFB, Ohio, 1960.

Miller, R. B. Analysis and specification of behavior for training. In R. Glaser (Ed.) Training research and education. Pittsburgh: University of Pittsburgh Press, 1962. Pp. 31-62.

Mirabella, A., & Lamb, J. C. Computer-based adaptive training applied to symbolic displays. NAVTRADEVCEN-TR 1594-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1966.

Mohler, S. R. Fatigue in aviation activities. AM 65-13, Federal Aviation Agency, Civil Aeromedical Research Institute, Oklahoma City, Okla., 1965.

Montague, W. E., Webber, C. E., & Adams, J. A. The effects of signal and response complexity of eighteen hours of visual monitoring. Human Factors, 1965, 7, 163-172.

Moore, H. G. The effects of load and accessibility of information upon performance of small teams. Contract AF 49(638)-449, U.S. Air Force Office of Scientific Research, Washington, D.C., Oct. 1961.

Morris, Ailene, & Harne, E. P. (Eds.) Visual search techniques: proceedings of a symposium. Publication 712, National Research Council, Washington, D.C., 1960.

Mowbray, G. H., & Gebhard, J. W. Man's senses as information channels. CM 936, Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Md., 1958.

Mullins, C. J., & Cox, J. A. Evaluation of the AFROTC flight instruction program. WADD-TN 60-44, Wright Air Development Division, Personnel Laboratory, Lackland AFB, Texas, 1960.

Nagay, J. Revisions of the standard flight-check for the airline transport rating based on the airline tryout. Report No. 89, CAA Division of Research, CAA, Washington, D.C., 1949.

Nagay, J. A study of the semi-annual instrument check for airline pilots. Report No. 93, CAA Division of Research, CAA, Washington, D.C. 1950.

National Research Council Committee on Aviation Psychology. Flight check manual. CAA Division of Research, CAA, Washington, D.C., 1950.

National Research Council Committee on Selection and Training of Aircraft Pilots. An historical introduction to aviation psychology. Report No. 4, CAA Division of Research, CAA, Washington, D.C., 1942.

National Research Council Committee on Selection and Training of Aircraft Pilots. History and development of the Ohio State flight inventory. Part II: Recent versions and current applications. Report No. 51, CAA Division of Research, CAA, Washington, D.C., 1945.

Naylor, J. C. Parameters affecting the relative efficiency of part and whole training methods: a review of the literature. NAVTRADEVCEN-TR 950-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1962.

Naylor, J. C., & Briggs, G. E. Long-term retention of learned skills: a review of the literature. ASD-TR 61-390, Air Force Systems Command, Wright-Patterson AFB, Ohio, Aug. 1961.

Naylor, J. C., & Briggs, G. E. Team training effectiveness under various conditions. J. appl. Psychol., 1965, 49, 223-229.

Newell, F. D. Criteria for acceptable representation of airplane dynamic responses in simulators used for pilot training. NAVTRADEVCEN-TR 1146-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1962.

Nielsen, G. Studies in self-confrontation. Copenhagen, Denmark: Monksgaard, 1962.

Ornstein, G. N., Nichols, I. A., & Flexman, R. E. Evaluation of a contact flight simulator when used in an Air Force primary pilot training program. Part II: Effectiveness of training on component skills. AFPTRC-TR 54-110, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1954.

Osgood, C. E. Method and theory in experimental psychology. New York: Oxford Press, 1953.

Parker, J. F., Jr., & Downs, Judith E. Selection of training media. ASD-TR 61-473, Aeronautical Systems Division, Aerospace Medical Laboratory, Wright-Patterson AFB, Ohio, 1961.

Parker, J. F., Jr., & Fleishman, E. A. Prediction of advanced levels of proficiency in a complex tracking task. USAF, Wright Air Development Center, Report 59-255, Wright-Patterson AFB, Ohio, Dec. 1959.

Parker, J. F., Jr., & Fleishman, E. A. Use of analytical information concerning task requirements to increase the effectiveness of skill training. J. appl. Psychol., 1961, 45, 295-302.

Payne, T. A., Dougherty, Dora J., Hasler, S. G., Skeen, J. R., Brown, E. L., & Williams, A. C., Jr. Improving landing performance using a contact landing trainer. SPECDEVCEN 71-16-11, USN, Office of Naval Research, Special Devices Center, Port Washington, N.Y., 1954.

Pearson, R. G. Task proficiency and feelings of fatigue. SAM-TR 57-77. USAF School of Aviation Medicine, Randolph AFB, Texas, 1957.

Persselin, L. E. Programmed instruction technology for ballistic weapon systems. BSD-TDR 64-72, Air Force Systems Command, Ballistic Systems Division, Norton AFB, California, 1964.

Pfeiffer, M. G., Clark, W. C., & Danaher, J. W. The pilot's visual task. A study of visual display requirements. NAVTRADEVCEN-TR 783-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1963.

Poe, A. C., Jr., Jolley, O. B., & Prophet, W. W. INTACT: Integrated instrument contact primary flight training. Army Aviation Digest, July 1960, 10-11.

Pomarolli, R. S. The effectiveness of the naval air basic instrument trainer. Special Report 65-7, U.S. Naval Aviation Medical Center, Pensacola, Florida, 1965.

Pressey, S. L. Teaching machine (and learning theory) crisis. J. appl. Psychol., 1963, 47, 1-6.

Rathert, G. A., Creer, B. Y., & Douvillier, J. G. Jr. Use of flight simulators for pilot-control problems. NASA Memorandum 3-6-59A. National Aeronautics and Space Administration, Ames Research Center, Moffett Field, Calif., 1959.

Rathert, G. A., Creer, B. Y., & Sadoff, M. The use of piloted flight simulators in general research. AGARD Report 365, NATO Advisory Group for Aeronautical Research and Development, Paris, France, 1961.

Rawson, H. E. Flight simulator study of human performance during low-altitude high-speed flight. TRECOM-TR 63-52, U.S. Army Transportation Research Command, Ft. Eustis, Va., Nov. 1963.

Ray, J. T., Martin, O.E., Jr., & Alluisi, E. A. Human performance as a function of the work-rest cycle: a review of selected studies. Publication 882, National Research Council, Washington, D.C., 1961.

Reynolds, B., & Bilodeau, I. M. Acquisition and retention of three psychomotor tests as a function of distribution of practice during acquisition. J. appl. Psychol., 1952, 44, 19-26.

Ritchie, M. L., & Hanes, L. F. An experimental analysis of transfer effects between contact and instrument flight training. Report prepared by Ritchie, Inc., Dayton, Ohio, under Contract No. FA-WA-4691, Federal Aviation Agency, Washington, D.C., 1964.

Ritchie, M. L., & Michael, A. L. Transfer between instrument and contact flight training. J. appl. Psychol., 1955, 39 (3), 145-149.

Roby, T. B., & Lanzetta, J. T. An investigation of task performance as a function of certain aspects of work group-structure. TN 57-124, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1956.

Roby, T. B., & Lanzetta, J. T. Conflicting principles in man-machine system design. J. appl. Psychol., 1957, 41, 170-178.

Roman, J. Risk and responsibility as factors affecting heart rate in test pilots. The flight research program II. Aerosp. Med., 1965, 36 (6), 518-523.

Rosenberg, S., & Hall, R. The effects of different social feedback conditions upon performance in dyadic terms. J. abnorm. soc. Psychol., 1958, 57, 271-277.

Ruocco, J. N., Vitale, P. A., & Benfari, R. C. Kinetic cueing in simulated carrier approaches. NAVTRADEVCEN-TR 1432-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1965. (a)

Ruocco, J. N., Vitale, P. A., & Benfari, R. C. Kinetic cueing in simulated carrier approaches: Supplement I. Study details. NAVTRADEVCEN-TR 1432-1-S1, U.S. Naval Training Device Center, Port Washington, N.Y., 1965. (b)

Sadoff, M., & Wempe, T. E. Pilot-vehicle system simulation for low-altitude, high-speed flight. NASA-TN D-2793, National Aeronautics and Space Administration, Ames Research Center, Moffett Field, Calif., June 1965.

Schramm, W. The research on programmed instruction. An annotated bibliography. Bulletin No. 35, Office of Education, Department of Health, Education and Welfare, Washington, D.C., 1964.

Seale, L. M. Relative flight proficiency of SNJ trained students and T34-T28-SNJ trained students: performance of field carrier landing practice and carrier qualifications. Special Report No. 56-30, U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1956.

Sells, S. B. Human flight behavior in groups. Review 6-58, USAF School of Aviation Medicine, Randolph AFB, Texas, July 1958.

Senders, J. W. Human tracking behavior. Report U-ED6141, Minneapolis-Honeywell Regulator Co., Aeronautical Division, Minneapolis, Minn., Nov. 1959. (a)

Senders, J. W. Survey of human dynamics data and a sample application. Report 59-712, USAF Wright Air Development Center, Wright-Patterson AFB, Ohio, Nov. 1959. (b)

Shoenberger, R. W., Wherry, R. J., Jr., & Berkshire, J. R. Predicting success in aviation training. Project MR005.13-3003, Report No. 7, U.S. Naval School of Aviation Medicine, Pensacola, Florida, Sep. 1963.

Siskel, M., Lane, F. D., Powe, W. E., & Flexman, R. E. Intracrew communication of B-52 and KC-135 student and combat crews during selected mission segments. AMRL-TR 65-18, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1965.

Smith, B. J. Task analysis methods compared for application to training equipment development. NAVTRADEVCEN 1218-5, U.S. Naval Training Device Center, Port Washington, N.Y., 1965.

Smith, J. F., & Flexman, R. E. An instructional manual for using performance record sheets designed for primary pilot training. Human Resources Research Center, Lackland AFB, Texas, 1952.

Smith, J. F., Flexman, R. E., & Houston, R. C. Development of an objective method of recording flight performance. Technical Report 52-15, Human Resources Research Center, Lackland AFB, Texas, Dec. 1952.

Smith, R. G., Jr. The development of training objectives. Research Bulletin No. 11, Human Resources Research Office, George Washington University, Alexandria, Va., June 1964.

Smith, R. G., Jr. Controlling the quality of training. Technical Report 65-6, Human Resources Research Office, George Washington University, Alexandria, Va., June 1965.

Smode, A. F. Learning and performance in a tracking task under two levels of achievement information feedback. J. exp. Psychol., 1958, 56 (4), 297-304.

Smode, A. F., Beam, J. C., & Dunlap, J. W. Motor habit interference: a resume of the literature and the development of principles for its minimization in training. Contract Nonr-2515(00), Dunlap and Associates, Inc., Darien, Conn., Jan. 1959.

Smode, A. F., Gruber, A., & Ely, J. H. The measurement of advanced flight vehicle crew proficiency in synthetic ground environments. AMRL-TDR 62-2, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1962.

Smode, A. F., Gruber, A., & Ely, J. H. Human factors technology in the design of simulators for operator training. NAVTRADEVCEN-1103-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1963.

Smode, A. F., & Meyer, D. E. Research data and information relevant to pilot training: I. General features of Air Force pilot training and some research issues. AMRL-TR-66-99(I), Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, July 1966.

Smode, A. F., Post, T. J., & Meyer, D. E. Research data and information relevant to pilot training: II. Description of pilot job requirements and training practices for representative missions and associated aircraft. AMRL-TR-66-99(II), Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, July 1966.

Smode, A. F., Vallerie, L. L., & Kelley, C. R. Performance evaluation on the Sikorsky Flight Simulator. Report No. 2, Dunlap and Associates, Inc., Darien, Conn., Dec. 1962.

Snyder, H. L. Visual aspects of low level flight. In J. W. Miller (Ed.) Visual, display and control problems related to flight at low altitude. ONR Symposium Report ACR-95, Office of Naval Research, Washington, D.C., Mar. 1964. Pp. 27-37.

Snyder, M. B. Methods of recording and reporting task analysis information. In Uses of task analysis in deriving training and training equipment requirements. WADD-TR 60-593, Wright Air Development Division, Wright Patterson AFB, Ohio, Dec. 1960. Pp. 11-32.

Soliday, S. M. Effects of task loading on pilot performance during simulated low-altitude high-speed flight. TRECOM-TR-64-69, U.S. Army Transportation Research Command, Ft. Eustis, Va., Feb. 1965.

Soliday, S. M., & Schohan, B. A simulator investigation of pilot performance during extended periods of low-altitude, high-speed flight. NASA-CR 63, National Aeronautics and Space Administration, Washington, D.C., June 1964.

Stark, E. A., Wolff, P. C., & Haggard, D. F. Target detection: Study I. A preliminary investigation of the trainability of target detection and distance estimation skills. HumRRO Research Memorandum, Human Resources Research Office, U.S. Army Armor Human Research Unit, Ft. Knox, Ky., July 1961.

Stoller, F. H. Closed circuit television and video tape for group psychotherapy with chronic mental patients. Amer. Psychologist, 1964, 5, 342-343. (Abstract)

Stolurow, L. M. Essential principles of programmed instruction. TR 8, Training Research Laboratory, Department of Psychology, University of Illinois, Urbana, Ill., 1965. (a)

Stolurow, L. M. Computer-based instruction. TR 9, Training Research Laboratory, University of Illinois, Urbana, Ill., 1965. (b)

Strategic Air Command. Pilot standardization/evaluation grading. SAC flying training manual. SACM 51-4, Vol. 4, HQ, SAC, Offutt AFB, Omaha, Neb., 1966.

Summers, L. G., & Ziedman, K. A study of manual control methodology with annotated bibliography. NASA Contractor Report CR-125, National Aeronautics and Space Administration, Washington, D.C., Nov. 1964.

Sutter, E. L., Townsend, J. C., & Ornstein, G. N. The light plane as a pre-primary selection and training device: II. Analysis of training data. AFPTRC-TR 54-35, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1954.

Swanson, A. M. Notes on simulator instrumentation for measurement of pilot proficiency. AFPTRC-TM 57-3, Air Force Personnel and Training Research Center, Lackland AFB, Texas, May 1957.

Swets, J. A. Learning to identify non-verbal sounds: an application of a computer as a teaching machine. NAVTRADEVCEN-TR 789-1, U.S. Naval Training Device Center, Port Washington, N.Y., 1962.

Swets, J. A., Harris, J. R., McElroy, Linda S., & Rudloe, H. Further experiments on computer-aided learning of sound identification. NAVTRADEVCEN-TR 789-2, U.S. Naval Training Device Center, Port Washington, N. Y., 1964.

Sylvania Electric Products, Inc. The UDOFT flight simulation system. AMRL-TDR 63-133, 6570th Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, 1963.

Taylor, F. V., & Birmingham, H. P. That confounded system performance measure--a demonstration. Psychol. Rev., 1959, 66, 178-182.

Thomas, F. H. Low altitude aerial observation: an experimental course of instruction. Technical Report 80, Human Resources Research Office, Alexandria, Va., Oct. 1962.

Thomas, F. H. Aviator performance in the light weapons helicopter during nap-of-the-earth flight. Tenth Annual Army Human Factors Research and Development Conference, Ft. Rucker, Ala., 5-8 October 1964.

Thomas, F. H., & Caro, P. W., Jr. Training research on low altitude visual aerial observation: a description of five field experiments. Research Memorandum 8, Human Resources Research Office, U.S. Army Aviation Human Research Unit, Ft. Rucker, Ala., July 1962.

Townsend, J. C. Evaluation of the Link ME-1 basic instrument flight trainer. TN 56-84, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1956.

Townsend, J. C., & Flexman, R. Suggested ways of improving instruction in the primary pilot training program. AFPTRC-TR 54-126, Air Force Personnel and Training Research Center, Lackland AFB, Texas, 1954.

Townsley, D. The philosophy of full mission training. WADC-TN 59-367, Wright Air Development Division, Air Research and Development Command, Wright-Patterson AFB, Ohio, 1960.

U.S. Air Force handbook of instructions for aerospace personnel subsystem designers. AFSCM 80-3, Headquarters Air Force Systems Command, Andrews Air Force Base, Md., July 1962.

University of Illinois Institute of Aviation. Simultaneous contact-instrument flight training. University of Illinois Institute of Aviation Aeronautics Bulletin No. 18, Jan. 1956.

Van Cott, H. P. Part versus whole practice in training: a discussion and review of the literature. American Institute for Research, Pittsburgh, Pa., 1955.

Vanderplas, J. M. Transfer of training and its relation to perceptual learning and recognition. Psychol. Rev., 1958, 65, 375-385.

Viteles, M. The aircraft pilot: five years of research, a summary of outcomes. Report No. 46, CAA Division of Research, CAA, Washington, D.C., 1945.

Viteles, M., & Backstrom, O. An analysis of graphic records of pilot performance obtained by means of the r-s ride recorder. Part I. Report No. 23, CAA Division of Research, CAA, Washington, D.C., 1943.

Viteles, M., & Thompson, A. The use of standard flights and motion photography in the analysis of aircraft pilot performance. Report No. 15, CAA Division of Research, CAA, Washington, D.C., 1943.

Viteles, M., & Thompson, A. An analysis of photographic records of aircraft pilot performance. Report No. 31, CAA Division of Research, CAA, Washington, D.C., 1944.

Voas, R. B., Bair, J. T., & Ambler, Rosalie K. The relationship between behavior in a stress situation and later separation from flight training with expressed anxiety toward flying. Research Report No. NM 001-101-108.01, Naval School of Aviation Medicine, Pensacola, Florida, June 1955.

Wade, J. E. Effectiveness of training procedures for low-altitude target acquisition and recognition. APGC-TR 64-68, Air Proving Ground Center, Eglin AFB, Florida, Oct. 1964.

Waller, T. G., & Wright, R. H. The effect of training on accuracy of angle estimation. HumRRO-TR 65-8, Human Resources Research Office, Alexandria, Va., 1965.

Walston, C. E., & Warren, C. E. A mathematical analysis of the human operator in a closed-loop control system. Report TR 54-96, USAF Personnel and Training Research Center, Lackland AFB, Texas, Dec. 1954.

Wapner, S., Odbert, H., & Festinger, L. Comparison of student pilot performance in successive check flights as measured by photographic records. Report No. 59, CAA Division of Research, CAA, Washington, D.C., 1946.

Webb, W. B. Comments reported in P. H. DuBois and W. H. Manning (Eds.) Methods of research in technical training. Report of conference held at U.S. Naval Air Station, Memphis, Tenn., Oct. 1956.

Webber, C. E. Survey study of the use of analog-digital conversion techniques and high speed digital computer for the treatment of pilot performance in a flight simulator. Report 58-1, University of Illinois, Aviation Psychology Laboratory, Jan. 1958.

Weisz, A. Z., & McElroy, Linda S. Response and feedback techniques for automated training of visual identification skills. NAVTRADEVCEN-TR 789-3, U.S. Naval Training Device Center, Port Washington, N.Y., 1964.

Wherry, R. J., Jr., & Curran, P. M. A study of some determiners of psychological stress. NSAM-941, U.S. Naval Aviation Medical Center, Pensacola, Florida, July 1965.

Whittenburg, J. A., & Wise, J. E. Feasibility for research application of visual attachments for dynamic flight simulators. Report No. 2: Requirements and state-of-the-art evaluation. HSR-RR 62/10-MK-X, Human Sciences Research, Arlington, Va., for the Federal Aviation Agency, Washington, D.C., 1963.

Wilcoxon, H. C., Davy, E., & Webster, J. C. Evaluation of the SNJ operational flight trainer. SPECDEVCEN-TR 999-2-2, Office of Naval Research, Special Devices Center, Port Washington, N.Y., 1954.

Wilcoxon, H. C., Johnson, W., & Golan, D. L. The development and tryout of objective check flights in pre-solo and basic instrument stages of naval air training. U.S. Naval School of Aviation Medicine, Pensacola, Florida, 1952.

Williams, A. C., Jr., & Flexman, R. E. Evaluation of the School Link as an aid in primary flight instruction. University of Illinois Institute of Aviation Aeronautics Bulletin No. 5, 1949.

Williams, A. C., Jr., MacMillan, J., & Jenkins, J. Preliminary experimental investigations of "tension" as a determinant of performance in flight training. Report No. 54, CAA Division of Research, CAA, Washington, D.C., 1946.

Willis, M. P. Deriving training device implications from learning theory principles. NAVTRADEVCEN 784-1, U.S. Naval Training Device Center, Port Washington, N.Y., July 1961.

Wolff, P. C. Target detection: Study 2. The comparative effects of continuous and delayed point-out during training in target detection. HumRRO Research Memorandum, Human Resources Research Office, Alexandria, Va., March 1961.

Wolff, P. C. Target detection: Study 7. Partial point-out of targets as collective reinforcement in group target detection training. HumRRO Research Memorandum, Human Resources Research Office, Alexandria, Va., Aug. 1962.

Wolff, P. C., Burnstein, D. D., & Van Loo, J. A. Target detection: Study 6. The effects of schedules of collective reinforcement on a class during training in target detection. HumRRO Research Memorandum, Human Resources Research Office, Alexandria, Va., July 1962.

Wolff, P. C., & Van Loo, J. A. Target detection: Study 3. The relative usefulness of active participation and verbal description techniques in target detection training. HumRRO Research Memorandum, Human Resources Research Office, Alexandria, Va., July 1962.

Woodworth, R. S., & Schlosberg, H. Experimental psychology (Rev. ed.) New York: Holt, Rinehart, & Winston, 1954.

Wright Air Development Division. Uses of task analysis in deriving training and training equipment requirements. WADD-TR 60-593, Wright-Patterson AFB, Ohio, 1960.

Zavala, A., et al. The analysis of helicopter pilot performance. Contract No. DA 49-193 MD 2632, Army Surgeon General's Office, Washington, D.C., June 1965.

Zeller, A. F., & Burke, J. M. Accidents of instructor pilots. An evaluation of instructor pilot accident rates in relation to total and and total instructor pilot time. Report 28-56, Directorate of Flight Safety Research, Norton AFB, Calif., 1956.

BLANK PAGE

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) BioTechnology, Inc. Arlington, Virginia 22210	2a REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b GROUP N/A

3 REPORT TITLE

AN ASSESSMENT OF RESEARCH RELEVANT TO PILOT TRAINING

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final report 1 February 1966 - 30 June 1966

5 AUTHOR(S) (Last name, first name, initial)

Smode, Alfred F.
Hall, Eugene R.
Meyer, Donald E.

6 REPORT DATE	7a TOTAL NO. OF PAGES	7b NO OF REFS
NOVEMBER 1966	241	336

8a CONTRACT OR GRANT NO. AF33(615)-2968 b PROJECT NO 1710 c Task no. 171003 d	9a ORIGINATOR'S REPORT NUMBER(S) 9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AMRL-TR-66-196

10 AVAILABILITY/LIMITATION NOTICES This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of Aerospace Medical Research Laboratories (MRHTO) Wright-Patterson Air Force Base, Ohio 45433

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY Aerospace Medical Research Laboratories Aerospace Medical Division, Air Force Systems Command, Wright-Patterson AFB, Ohio

13 ABSTRACT

This report presents a critical review and interpretation of the considerable amount of research data that have either direct or indirect implications for the training of pilots. The purpose is to organize systematically the research findings from the human performance and the training research literature that are pertinent to pilot training, and, based on the status of research in defined areas, to identify researchable issues. Successive portions of the report deal with studies on the measurement simulation and transfer of training, operational components of the pilot's job, and the maintenance of flying proficiency. In addition, attention is given to studies concerned with improving training systems and recent innovations in training methods are reviewed. As it provides a considerable background of information directly concerned with pilot training, this report will be of interest to individuals involved in any aspect of flight training.

DD FORM 1 JAN 64 1473

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Military training Pilot training programs Pilot skill requirements Training devices Flying curricula						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ______________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ______________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ______________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

SUPPLEMENTARY

INFORMATION

IDENTIFICATION	FORMER STATEMENT	NEW STATEMENT	AUTHORITY
AD-804 600 Biotechnology Inc., Arlington, Va. Final rept. 1 Feb-30 Jun 66. Rept. no. AMRL-TR-66-196 Nov 66 Contract AF 33(615)-2968	No Foreign without approval of Aerospace Medical Research Labs., Attn: MRHTO, Wright-Patterson AFB, Ohio.	No Foreign without approval of Air Force Human Resources Laboratory, Attn: HRT, Wright-Patterson AFB, Ohio.	AMRL, USAF ltr, 22 Dec 69